SW

Vol. 34, No. 3 PROPERTY OF PRINCETON UNIVERSITY LIBRARY RECEIVED JUL 26 1944 June, 1944

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, ON LEAVE
FRANCIS W. IRWIN, ACTING EDITOR
UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CHARLES W. BRAY, *Princeton University;* ELMER K. CULLER, *University of Rochester;* CLARENCE H. GRAHAM, *Brown University;* JOY P. GUILFORD, *University of Southern California;* FRANCIS W. IRWIN, *University of Pennsylvania;* DONALD G. MARQUIS, *Yale University;* ARTHUR W. MELTON, *University of Missouri;* CARROLL C. PRATT, *Rutgers University.*

CONTENTS

An Experimental Study of the Drawing Behavior of Adult Psychotics in Comparison with That of a Normal Control Group: A. ANASTASI AND J. P. FOLEY, JR. 169

An Experimental Study of the Role of the Ego in Work: II. The Significance of Task-Orientation in Work: H. B. LEWIS AND M. FRANKLIN 195

The Effect of Different Amounts of Reinforcement upon the Acquisition and Extinction of a Simple Running Response: F. A. MOTE 216

Guthrie's Theory of Learning: A Second Experiment: J. P. SEWARD, J. B. DILL, AND M. A. HOLLAND 227

The Problem of Stimulation Deafness: I. Cochlear Impairment as a Function of Tonal Frequency: E. G. WEVER AND K. R. SMITH 239

An Experimental and Statistical Study of Olfactory Preferences: M. D. EYSENCK 246

A Dual Selective Amplifier: H. B. CARLSON AND A. J. EBEL 253

PUBLISHED BI-MONTHLY,
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

XUM

PUBLICATIONS OF

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

WILLARD L. VALENTINE, *Business Manager*

PSYCHOLOGICAL REVIEW

HERBERT S. LANGFELD, *Editor*
Princeton University

Contains original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.
Subscription: $5.50 (Foreign, $5.75). Single copies, $1.00.

PSYCHOLOGICAL BULLETIN

JOHN E. ANDERSON, *Editor*
University of Minnesota

Contains critical reviews of books and articles, psychological news and notes, university notices, and announcements. Appears monthly (10 issues), the annual volume comprising about 665 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

S. W. FERNBERGER, *Editor, on Leave,* FRANCIS W. IRWIN, *Acting Editor*
University of Pennsylvania

Contains original contributions of an experimental character. Appears bi-monthly (beginning 1944), one volume per year, each volume of six numbers containing about 520 pages.
Subscription: $7.00 (Foreign, $7.25). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

WALTER S. HUNTER, *Editor*
Brown University

Appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.
Subscription: $7.00 (Foreign, $7.25). Single copies, 75c.

PSYCHOLOGICAL MONOGRAPHS

JOHN F. DASHIELL, *Editor*
University of North Carolina

Consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.
Subscription: $6.00 per volume (Foreign, $6.30).

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

GORDON W. ALLPORT, *Editor*
Harvard University

Appears quarterly, January, April, July, October, the four numbers comprising a volume of 560 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.
Subscription: $5.00 (Foreign, $5.25). Single copies, $1.50.

JOURNAL OF APPLIED PSYCHOLOGY

DONALD G. PATERSON, *Editor*
University of Minnesota

Covers the applications of psychology in business, industry, education, etc. Appears bi-monthly, February, April, June, August, October, and December.
Subscription: $6.00. Single copies, $1.25.

Subscriptions, orders, and business communications should be sent to

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

XUM

Journal of
Experimental Psychology

VOL. 34, NO. 3 JUNE, 1944

AN EXPERIMENTAL STUDY OF THE DRAWING BEHAVIOR OF ADULT PSYCHOTICS IN COMPARISON WITH THAT OF A NORMAL CONTROL GROUP [1]

BY ANNE ANASTASI

Queens College

AND

JOHN P. FOLEY, JR.

The Psychological Corporation

Drawings by psychotics have been explored from a variety of viewpoints: as artistic products, as psychiatric devices for diagnosis and therapy, and as sources of psychological information on the nature of such processes as perception, thinking, imagination, and æsthetic behavior. The study of such products has a long history, but has shown an especially rapid development in all its phases within the last decade. Relatively little of a conclusive nature can, however, be gleaned from the voluminous literature in this field, owing to the inadequate control of experimental conditions and to the excessive theoretical speculation which frequently obscures the data.

In the effort to arrive at a systematic analysis of the differentiating characteristics of psychotic drawings, the writers have followed three

[1] The present study is part of a project on the artistic behavior of the abnormal, conducted by the writers under the auspices of the Columbia University Council for Research in the Social Sciences. A detailed report of the entire project is to appear later in book form.

The writers wish to acknowledge the coöperation of the staffs of the hospitals at which data were obtained and to express special appreciation to the superintendents of such hospitals, listed as follows in order of number of cases supplied: Dr. Winfred Overholser, Saint Elizabeths Hospital, Washington, D. C.; Dr. Karl Bowman, Psychiatric Division of Bellevue Hospital, New York City; Dr. Nolan D. C. Lewis, New York State Psychiatric Institute and Hospital, New York City; Dr. James W. Milligan, Madison State Hospital, North Madison, Indiana; and Dr. John M. Mras, Chester County Hospital for the Insane, Embreeville, Pennsylvania.

Acknowledgment is also made to Dr. N. Henry Pronko and Mr. John W. Molyneaux for their invaluable assistance in collecting the data.

XUM

lines of investigation. The first consisted of a survey of the literature (**1**, **2**, **3**, **4**) for the following types of material: printed reproductions of art products by the abnormal; case studies and biographies of artists who have been institutionalized as insane at some time in their lives; case studies and experimental investigations conducted on patients in mental hospitals; and theoretical discussions on the relationship between art and insanity and on the creative activities of the abnormal. The second approach was through an analysis of over 1200 drawings spontaneously produced by 212 patients institutionalized in 62 mental hospitals in this country (**5**). The third part of the project, of which the present paper is a summary, is an experimental investigation of the drawing behavior of institutionalized patients, under controlled conditions, in comparison with that of a matched control group of normal subjects.

Method

1. *Procedure and materials.*

Four drawings were obtained from each *S*, in the following sequence and with the verbal directions reproduced below:

(1) *Free choice:* "On *this* piece of paper, draw a picture of *anything* you want to—anything at all you want to draw, and color it to suit yourself."
(2) *Danger:* "This time make a drawing to show *danger*. Make a picture which means danger to you—which shows people what danger is."
(3) *Man:* "Now this will be easier. Draw a picture of a man." (A drawing of either a whole man or of head or bust was accepted, and the *S* was informed of this fact if he inquired.)
(4) *Copy of a stylized floral design in six colors* (model reproduced in Fig. 13): "This will be still easier. This time I just want you to *copy* this drawing. On this piece of paper make a drawing exactly like this one and color it the same way. I'll pick out the crayons you need."

Each *S* was tested individually. Uniform drawing paper, pencil, eraser, wax crayons,[2] and scrap paper were presented in a standard, predetermined arrangement. A detailed record was kept of the *S*'s comments and behavior during the drawing. Time records, obtained inconspicuously with a stop watch, included the *delay* which elapsed between the completion of the directions and the making of the first mark on paper by the *S*, as well as *total time* elapsing between the completion of the directions and the submitting of the drawing by the *S*.

2. *Subjects.*

A total of 680 white, adult *S*s were employed, including 340 patients institutionalized in five mental hospitals and 340 normal control *S*s. Each of these two groups consisted of 170 men and 170 women. The abnormal and control groups were equated [3] in age, marital status, educational and occupational level, artistic training and experience, geographical distribution (with reference to birthplace and major residence), and national background. These characteristics of the *S*s, together with other relevant information recorded on an individual data sheet, were obtained chiefly by direct questioning in the case of the normal *S*s and from the hospital records for the abnormal *S*s. Certain other supplementary and corroborative sources of information were consulted, whenever available, for both abnormal and normal *S*s.

[2] Prang Crayonex, American Pencil Company, No. 323, containing an assortment of 24 colors.

[3] The degree to which it proved possible to equate the two groups in these various characteristics can be determined by an examination of the distributions, means, SD's, and other relevant data in the more detailed report which is to appear in book form.

The clinical diagnoses of the abnormal *Ss* cover a wide variety of the traditional syndromes. The largest group (51 percent) are diagnosed as schizophrenia, all sub-types being represented. Manic-depressive psychosis, alcoholic psychoses, psychoneuroses, involutional melancholia, and paresis follow in decreasing order of frequency. A few other conditions are represented by very small groups. Evidence of intellectual deficiency as a complicating factor was present in only 11 cases. It should be noted that all of the diagnoses were obtained from the hospital records and are independent of the drawing behavior or other observations made in the course of the present experiment. Length of institutionalization ranged from one day to 24 years, averaging 17.61 months. Approximately three-fourths of the patients had been institutionalized for less than one year.

Results

1. *Subject-matter of 'free choice' drawings (Drawing No. 1).*

With respect to content, the 'free choice' drawings fall into 19 major categories. These categories, together with the percentage of drawings in abnormal and normal groups, respectively, in each category and the significance ratios of the differences between these two percentages, are given in Table I. Unless otherwise stated, all

TABLE I

Subject-Matter Classification of Drawing No. 1: 'Free Choice'

Subject-matter	Percentage Frequency		Significance Ratio of Abnormal-Normal Difference*
	Abnormal	Normal	
1. Person	19.12	14.71	1.54
2. Animal	5.00	9.12	−2.11
3. Flower	6.47	10.88	−2.04
4. Tree	2.06	5.29	−2.24
5. Fruit	1.47	2.65	−1.07
6. Building	13.24	14.12	−0.33
7. Landscape or seascape	9.41	11.47	−0.88
8. Transportation	6.18	8.24	−1.04
9. Household equipment	6.18	5.88	0.15
10. Graphic materials	2.06	1.47	0.58
11. Tools and instruments	1.76	0.00	2.44
12. Box or basket	4.41	3.24	0.80
13. Designs	2.65	2.65	0.00
14. Fantastic compositions	2.94	0.00	3.23
15. Unrelated objects	9.12	2.35	3.82
16. Scribbling or scrawl	1.76	0.00	2.44
17. Writing only	1.76	0.00	2.44
18. Miscellaneous	3.82	7.94	−2.30
19. Refusal to draw	0.59	0.00	1.40

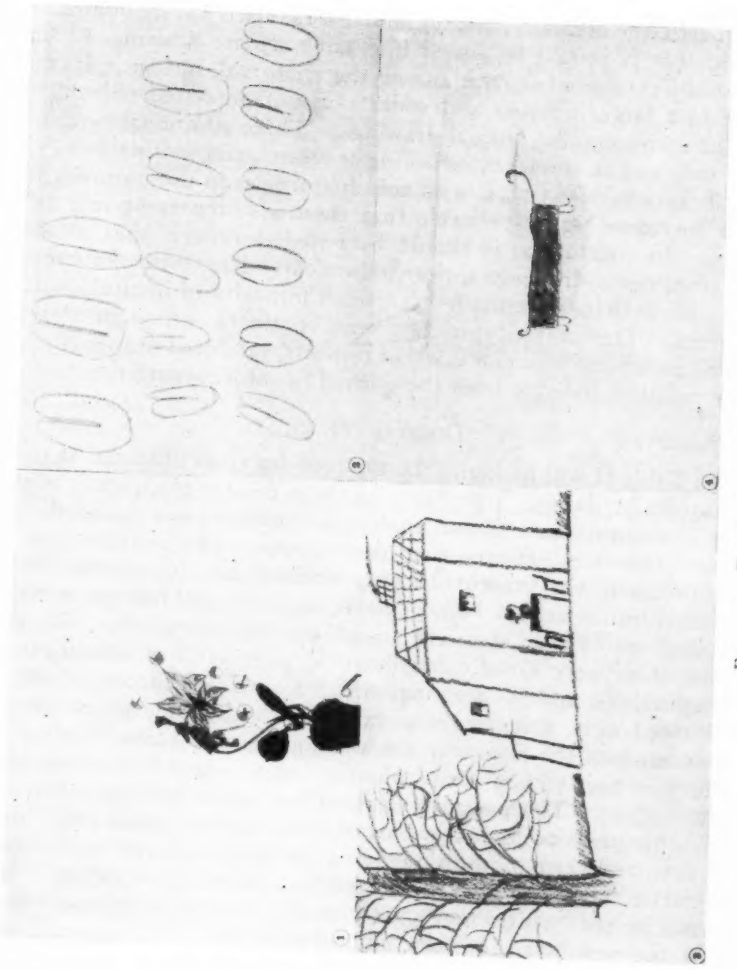
* In a composite chi-square analysis of the complete abnormal and normal distributions, $P < .01$, i.e., the chances that the entire pattern of abnormal-normal differences resulted from sampling error are practically zero.

categories are both singular and plural. Some include sub-categories. Thus, *transportation* includes ships, automobiles, airplanes, and bicycles; *household equipment* includes tableware and kitchen utensils as well as furniture and appurtenances; *graphic materials* cover reading matter and writing and drawing accessories; *designs* are sub-

divided into single geometric figures, such as circle, star, or triangle, as well as more complex or repetitive decorative designs. Under *fantastic compositions* are to be found maps and plans of paranoid 'inventions' as well as a variety of symbolical, abstract, or incomprehensible drawings. The latter often contain a welter of meaningless lines or forms, highly distorted objects, and idiosyncratic symbolism. Those drawings classified as unrelated objects portray discrete, disconnected, but recognizable objects, the number within any one drawing varying from two to 13 among the abnormal and from two to four among the normal drawings. *Scribbling or scrawl* and *writing* were submitted by a few abnormal *S*s in place of a drawing. Into these two categories were placed only those specimens which contained no drawing, but only writing or a single scrawl or scribbled lines. The *miscellaneous* category includes the following sub-divisions: personal objects and wearing apparel, diagrams and maps (exclusive of delusional projects), interiors, conventionalized symbols (such as flags or American eagle), bridges, religious and allegorical drawings, match or match-box, and eight other individual drawings not elsewhere classifiable. All of the items in the *miscellaneous* category were so classified because of the low frequency with which each occurred. Finally, under *refusal to draw* are to be found two abnormal *S*s who could not be induced to make drawing No. 1, although both agreed to copy the sample design of drawing No. 4.

An examination of Table I will show that there is extensive *overlapping* in the type of subject-matter portrayed by abnormal and normal groups in the 'free choice' drawing. Most of the subject-matter categories do not differentiate significantly between the abnormal and normal drawings. Certain differences may, however, be noted between the two groups. In the first place, the sheer *range of subject-matter* covered by the abnormal drawings is greater than that of the normal drawings. Excluding for this purpose *refusal to draw*, we find abnormal entries in all of the 18 subject-matter categories, whereas no normal entries occur in four categories. Further corroboration is furnished if the sub-categories are included in the analysis. A combined total of seven categories and sub-categories are found to contain no normal drawings, while abnormal entries are absent from only two sub-categories. This greater diversity of subject-matter in the abnormal drawings is to be expected from individuals much of whose behavior is aberrant or idiosyncratic.

Certain differences within *specific subject-matter categories* should also be noted. A reliably greater percentage of abnormal than of normal drawings are classified as *fantastic compositions* and *unrelated objects*, in both cases the difference in percentage being over three times as large as its standard error. The former category, further-



FIGS. 1–4. Drawing No. 1: 'Free Choice'

more, contains no normal drawings, and may therefore be regarded as highly indicative of abnormality. Such fantastic compositions usually reflect the delusional, disintegrated, and other idiosyncratic behavior of the patient. The drawings of unrelated objects further emphasize the diversity, scatter, and disorganized nature of abnormal products. It should be added that not only are drawings of unrelated objects more common among the abnormal, but such drawings contain a larger number and wider range of unrelated objects than do the corresponding normal drawings. Three additional categories, *viz.*, *tools and instruments*, *scribbling or scrawl*, and *writing only*, yield significance ratios of 2.44, with zero frequencies in the normal group. It is therefore highly probable that these also represent real differences. In conclusion, it should be noted, however, that although certain types of drawings appear to be clearly diagnostic of abnormality, such drawings are made by a small minority of institutionalized patients. The majority of abnormal patients, although showing clearly psychotic behavior in other respects, produced drawings which were indistinguishable from the normal in subject-matter.

2. *Portrayals of 'danger' (Drawing No. 2).*

In Table II will be found the types of drawings obtained as representations of 'danger,' together with the percentage frequency of each type in abnormal and normal groups, respectively, and the reliabilities of the differences between such percentages. The restricted nature of the task in the present drawing made it possible to classify the drawings into relatively homogeneous categories, all but two of which required no further differentiation into sub-categories. *Fantastic compositions* were sub-divided into: (*a*) portrayals of delusions and hallucinations, and (*b*) drawings which were either incomprehensible or showed only a remote association with the concept of danger. An example of the former is a drawing of a ward scene showing the patient in bed, with a white woman and a Negro woman standing near the bed. The patient's explanation of the element of danger in this picture was made in terms of her delusional belief that "bats, hell cats, and hermaphroditic women" in her ward rape her at night. The latter sub-category is illustrated by a drawing consisting of four scrawls in red, white, green, and black, respectively, representing "holes in rocks, earthquakes, underground fires, sinking, explosion, eruption, poison gas, small pox, typhoid fever, black death, parrot fever." Into this category were placed all highly idiosyncratic symbolizations of danger. Drawings employing conventionalized symbols (e.g., swastika) or portraying danger through well-known allegories (e.g., ostrich burying its head in the sand), on the other hand, were classified under *abstract or conventionalized symbolism.*

XUM

TABLE II

SUBJECT-MATTER CLASSIFICATION OF DRAWING NO. 2: 'DANGER'

Subject-matter	Percentage Frequency		Significance Ratio of Abnormal-Normal Difference*
	Abnormal	Normal	
1. Traffic	30.88	36.47	−1.54
2. Conventional sign or signal	10.59	10.88	−0.12
3. Skating, ice	1.47	2.94	−1.30
4. Falling	6.18	5.00	0.67
5. Drowning, sinking, flood	2.35	2.94	−0.48
6. Falling objects, explosion	1.76	1.76	0.00
7. Arms and explosives	2.94	6.47	−2.07
8. War	0.88	2.06	−1.23
9. Fire	7.65	10.88	−1.45
10. Lightning, electricity	1.76	4.41	−2.01
11. Animals	0.88	4.71	−2.99
12. Abstract or conventionalized symbolism	0.88	2.35	−1.46
13. Fantastic compositions	7.65	0.00	5.28
14. Several discrete objects	2.65	0.59	2.12
15. Scribbling or scrawl	1.76	0.00	2.44
16. Writing only	3.53	0.29	3.12
17. Miscellaneous	6.76	7.65	−0.45
18. Recognizable object not representing danger	2.35	0.29	2.34
19. Refusal to draw	4.12	0.29	3.42
20. No data for other reasons	2.94	0.00	3.23

* In a composite chi-square analysis of the complete abnormal and normal distributions, $P < .01$, i.e., the chances that the entire pattern of abnormal-normal differences resulted from sampling error are practically zero.

This category is in turn to be distinguished from that covering *conventional sign or signal*, which refers to the reproduction of objects commonly employed to designate danger, such as red flags, railroad crossing signs, and the skull-and-crossbones label for a poison bottle.

The *miscellaneous* category contains the following sub-categories: sex, alcohol (cf. Fig. 5—"man drinking"), poison, illness, hanging, storms, physical restraint (prison door, barred hospital window), social and economic dangers (unemployment, Nazism, etc.), person's reaction to danger, and a few individual cases not elsewhere classifiable. *Recognizable object not representing danger* refers to any comprehensible drawing which is clearly unrelated to danger. Finally, it will be noted that to the category *refusal to draw*, found in drawing No. 1, has been added a further category designated as '*no data for other reasons*.' Into the latter category were placed cases in which the directions for drawing No. 2 were not even given, owing to the *S*'s previously shown lack of coöperation or to his being prevented from drawing by other circumstances beyond his control.

As in the case of drawing No. 1, it will be noted that most of the categories fail to differentiate significantly between abnormal and normal drawings. Approximately one-third of both abnormal and

FIGS. 5–6. Drawing No. 2: 'Danger'
FIGS. 7–8. Drawing No. 3: 'Man'

normal Ss represented danger by some *traffic* situation, involving automobiles, pedestrians, intersections, railroads, and the like. To this should be added most of the 10 or 11 percent in each group who drew a *conventional sign or signal*, traffic signals being the most frequent item in this category.

Four categories yield significance ratios larger than 3.00 between the abnormal and normal percentages. The largest difference, with a significance ratio of 5.28, is obtained in the case of *fantastic compositions*, which were produced by 7.65 percent of the abnormal and no normal Ss. *Writing only*, *refusal to draw*, and *no data for other reasons* all yield significance ratios larger than 3.00. To these may be added three other categories, *viz.*, *several discrete objects*, *scribbling or scrawl* (cf. Fig. 6), and *recognizable object not representing danger*, all of which have higher frequencies in the abnormal group, with significance ratios larger than 2.00. It will be noted that all of the categories which contain a significant excess of abnormal drawings indicate relatively idiosyncratic, disorganized, and uncoöperative behavior in the drawing situation. As in the case of drawing No. 1, however, the total percentage of abnormal Ss whose drawings fell into such categories is small.

Another finding which corroborates the results obtained with drawing No. 1 is the greater scatter of subject-matter in the abnormal drawings, as indicated by the number of categories and sub-categories which contain entries from the abnormal and normal groups, respectively. Only one sub-category, social and economic dangers, contains no abnormal drawings. A combined total of six categories and sub-categories, on the other hand, are found exclusively in the abnormal group.

3. *Classification of drawings of 'man' (Drawing No. 3).*

The drawings of 'man' were analyzed in reference to a number of general characteristics of content and technique, which have been summarized in Table III. As in the two preceding drawings, the percentage frequency of each category in the abnormal and normal groups and the reliabilities of the differences between the two have been reported. It will be noted that a significantly larger percentage of abnormal Ss show *refusal to draw*,[4] the significance ratio of the abnormal-normal difference being 2.98. For 20 abnormal and one normal Ss, there are *no data for other reasons*,[4] the significance ratio of this difference being 4.23.

With respect to *amount of figure shown*, it is found that a reliably greater number of normal Ss draw only the *head or bust*, the difference

[4] These categories should be repeated within each drawing characteristic in Table III in order that the total percentage of drawings within each characteristic shall be 100.

XUM

TABLE III

CLASSIFICATION OF DRAWING NO. 3: 'MAN'

Drawing Characteristic	Percentage Frequency		Significance Ratio of Abnormal-Normal Difference*
	Abnormal	Normal	
Amount of figure shown:			
Head or bust	26.47	40.29	−3.86
Complete or three-quarters man	60.59	59.12	0.39
Cannot judge	3.82	0.00	3.67
View portrayed:			
Profile	41.47	55.00	−3.56
Full- or three-quarters-face	33.82	31.47	0.65
Both	11.47	12.94	−0.59
Cannot judge	4.12	0.00	3.81
Type of man drawn:			
General 'man'	78.24	84.41	−2.07
Particular man	9.71	15.00	−2.10
Cannot judge	2.94	0.00	3.23
Style of drawing:			
Schematic or stylized	17.94	15.00	1.04
Realistic	65.82	82.65	−5.12
Cannot judge	7.06	1.76	3.40
Fullness of detail:			
Slight detail, gross impression	30.59	20.88	2.91
Moderate detail	50.29	71.18	−5.71
Minute detail	7.06	7.35	−0.15
Cannot judge	2.94	0.00	3.23
Skillfulness of portrayal:			
Crude or poorly drawn	35.29	30.59	1.31
Mediocre technique and execution	44.12	57.65	−3.56
Very skillfully drawn	8.53	11.18	−1.16
Cannot judge	2.94	0.00	3.23
No drawing available:			
Refusal to draw	3.24	0.29	2.98
No data for other reasons	5.88	0.29	4.23

* Chi-square analyses of the abnormal and normal distributions yielded *P*-values smaller than .01 within each of the classifications reported in this table. Thus, in each case, the chances that the entire pattern of abnormal-normal differences resulted from sampling error are practically zero.

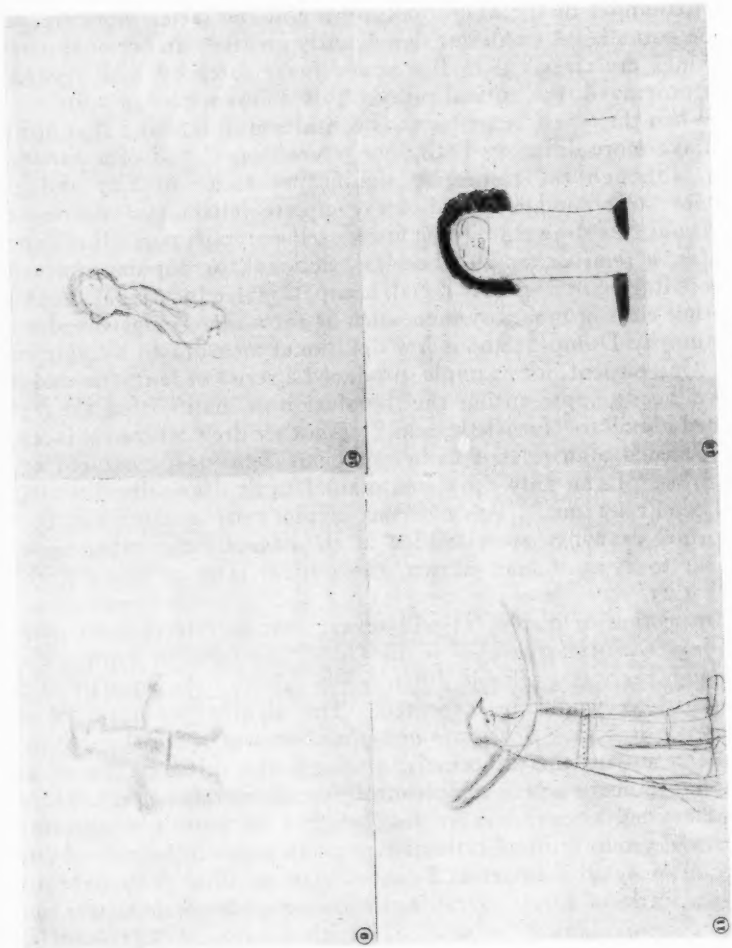
having a significance ratio of 3.86. This finding may reflect a greater awareness by the normal *Ss* of the limitations of their drawing ability. Drawings of a *complete or three-quarters man* show no significant difference in frequency between the two groups. A significantly greater number of abnormal drawings are found, however, in the *cannot judge* category (critical ratio of 3.67). The drawings so classified include both those in which no conceivable attempt was made to portray a man and those depicting a man but so incomplete, distorted, or crudely or indistinctly drawn as to render classification impossible. When all drawings are classified with respect to *view portrayed*, there is a significantly greater frequency of *profile* drawings by normal *Ss* (critical ratio of 3.56). Such a finding may also indicate a more widespread recognition of limitations in drawing ability

on the part of the normal *Ss*. There is not a statistically reliable difference between the two groups in the frequency of *full- or three-quarters-face* drawings or of drawings showing *both* profile and full-face or three-quarters-face views, although the former are slightly more common in the abnormal group and the latter more frequent in the normal. A small but significantly greater number of abnormal drawings are classified in the *cannot judge* category with regard to view portrayed, the critical ratio of this difference being 3.81.

When the *type of man drawn* is considered, it is found that normal *Ss* make more drawings both of a *general 'man'* and of a *particular man*, although the respective significance ratios of 2.07 and 2.10 are not large enough to indicate complete reliability. Among the *particular men* drawn are to be found: self-portrait; portrait of experimenter, a relative, or an historical, fictional, or popular character; representation of a specific racial, occupational, educational, religious, or other class of men; toy men, such as snow man, gingerbread man, or Humpty Dumpty; and a few delusional men drawn by abnormal *Ss*. One patient, for example, produced a series of fantastic and distorted heads, representing the "evolution of man" from the "pin-headed idiot" to "futuristic man." Another drew a bizarre, faceless, but bearded man related to her religious delusions, commenting as she drew, "I can only draw one man; I can't draw him like that—God won't let me." A small but significantly greater number of abnormal drawings are classified in the *cannot judge* category with respect to type of man drawn, the critical ratio of this difference being 3.23.

Inspection of Table III will reveal that in reference to *style of drawing*, *realistic portrayal* is the most common in both groups, although there is a reliable difference in favor of the normals in this category, as would be expected. The significance ratio of this difference is 5.12. *Schematic or stylized* drawings are slightly more frequent among the abnormals, although the critical ratio of 1.04 does not indicate a very significant difference in this respect. Drawings classified as *cannot judge* are rare, but are more common in the abnormal group (critical ratio of 3.40). In terms of *fullness of detail*, more drawings by abnormal *Ss* show *slight detail or gross impression* (critical ratio of 2.91). Drawings showing *moderate detail* are much more common among normal *Ss* (critical ratio of 5.71). *Minute detail* is found in relatively few drawings, and there is no observed difference between the two groups in this respect, contrary to much that has been claimed in the literature. Drawings classified in the *cannot judge* category with respect to fullness of detail are found rarely but with a significantly greater frequency among the abnormal *Ss*, the critical ratio of this difference being 3.23.

XUM



FIGS. 9–12. Drawing No. 3: 'Man'

Finally, with regard to *skillfulness of portrayal, crude or poorly drawn* products are slightly but not significantly more common in the abnormal groups (significance ratio of 1.31). Most of the drawings by both groups show *mediocre technique and execution*, with a significant difference favoring the normal group (significance ratio of 3.56), as would be expected. Very few products by either group are *very skillfully drawn*, there being a small and insignificant difference in this category in favor of the normal group (critical ratio of 1.16). Drawings classified in the *cannot judge* category with respect to skillfulness of portrayal occur rarely but with a significantly greater frequency among abnormal *S*s, the significance ratio of this difference being 3.23.

4. *Analysis of copies of sample design (Drawing No. 4).*

In Fig. 13 will be found a reduced reproduction of the stylized floral design which the *S* was directed to copy for drawing No. 4. The terms which we shall employ to designate the parts of this design, together with the crayons with which each part was colored, are as follows: stem and branch (grey); sepals (violet); petals (magenta); stamens (yellow-orange); leaf (yellow-green); leaf center (green). Reference will also be made to the pencilled 'loops,' two of which are found on the lower edge of the sepals and two on the lower edge of the petals; the 'scallops' outlining the sepals and petals; and the 'serrations' along the upper edge of the leaf.

TABLE IV

ANALYSIS OF DRAWING NO. 4: COPY

Classification of Copy	Percentage Frequency		Significance Ratio of Abnormal-Normal Difference*
	Abnormal	Normal	
1. Perfect or nearly perfect	7.35	10.59	−1.48
2. Moderately accurate	51.47	68.53	−4.61
3. Careless, crude	19.71	18.24	0.49
4. Barely recognizable	11.18	1.76	5.09
5. Incomplete	2.35	0.00	2.83
6. Scribbling or doodling	0.29	0.00	0.97
7. Writing only	0.29	0.00	0.97
8. Refusal to draw	0.29	0.00	0.97
9. No data for other reasons	7.06	0.88	4.18

* In a composite chi-square analysis of the complete abnormal and normal distributions, $P < .01$, i.e., the chances that the entire pattern of abnormal-normal differences resulted from sampling error are practically zero.

In Table IV, a general classification of the copies is presented, together with the percentage of abnormal and normal *S*s in each category and the reliability of the difference between the two. The

first four categories in the table may be considered together, since they represent varying degrees of proficiency in the execution of a complete copy of the design. It will be noted that the majority of drawings in both groups (51.47 percent abnormal and 68.53 percent normal) fall into the *moderately accurate* category. The percentage of such drawings among the normal, however, is reliably greater than that among the abnormal, the critical ratio of the difference being 4.61. Copies classified as *perfect or nearly perfect* are relatively few in both groups, with a slight though not reliable excess among the normal. *Careless, crude* copies occur with almost equal frequency in both groups. *Barely recognizable* reproductions, on the other hand, were made by 11.18 percent of the abnormal and 1.76 percent of the normal Ss, this difference yielding a significance ratio of 5.09. In this category are found the highly distorted, bizarre drawings characteristic of abnormality. Three such drawings are reproduced in Figs. 14, 15, and 16. It should be added that the few normal drawings in this category showed less extreme distortions than did most of the abnormal drawings. As in drawings No. 1, No. 2, and No. 3, however, it must be emphasized that such clearly 'abnormal' drawings are produced by only a small percentage of the entire abnormal sampling. The remaining categories in Table IV are self-explanatory and contain very few cases, although it is interesting to note that all are abnormal cases, with the exception of three normal men who through uncoöperativeness or for other reasons had discontinued the experiment before drawing No. 4 was reached.

The copies were also examined in reference to *omissions, additions, size errors, form errors,* and *coloring peculiarities.*[5] In 23 abnormal drawings, the categories of omissions, additions, and form errors were inapplicable, either because no attempt had been made to copy the given design or because the copy was too distorted to be judged in these respects. All drawings, however, permitted an analysis in terms of size and coloring.

With the exception of loop omissions, which were common in both groups and appear to be of little or no significance, all *omissions* were more common among the abnormal than among the normal. This discrepancy was increasingly apparent as the seriousness of the omission increased. Thus, omission of sepals, leaf, or branch was found only in the abnormal, whereas omission of the leaf center and all or some of the five stamens occurred in both groups, but with a greater frequency in the abnormal. *Additions* likewise show a fairly clear discrepancy between abnormal and normal drawings, qualitatively as well as quantitatively. All of the additions made by normal

[5] For the quantitative data pertaining to this analysis, the reader is referred to the more detailed report which is to appear in book form.

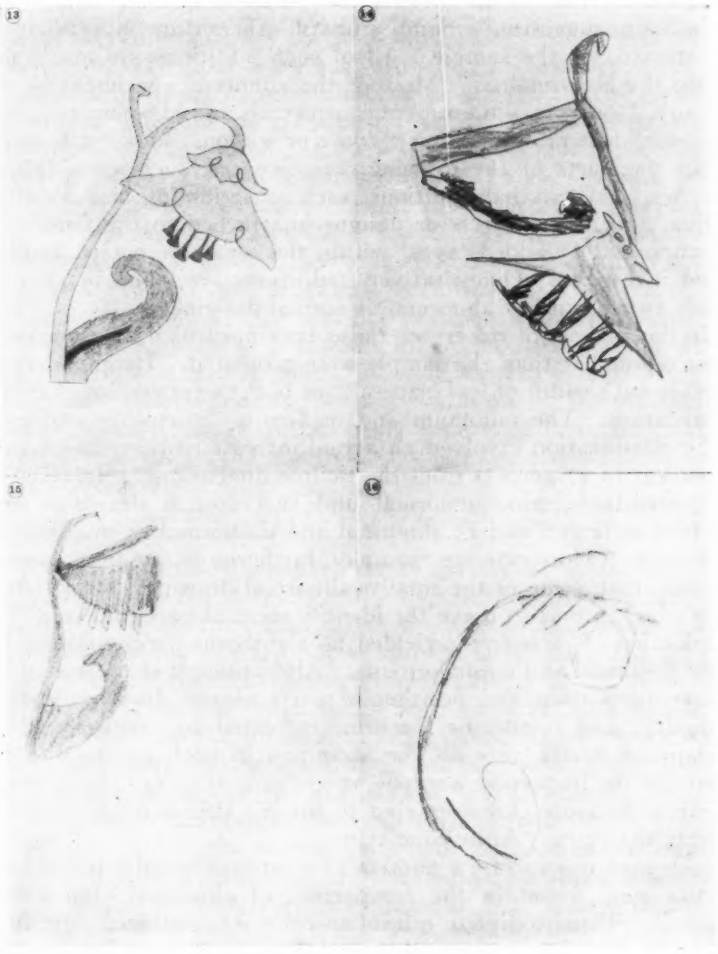


FIG. 13. Drawing No. 4: Sample Design
FIGS. 14–16. Drawing No. 4: Copies of Sample Design

*S*s are explicable in terms of drawing technique or careless observation of the sample. Some *S*s, for example, added a small oval at the tip of one or more stamens, which gave the stamen a solid appearance, as if viewed from below. Others added a horizontal line at the base of the stem, suggesting ground; and still others drew more than the five stamens in the sample. All of such additions are also found among the abnormal *S*s. Most of the abnormal specimens in this category, however, contain more conspicuous and idiosyncratic additions, such as a multiplication of loops or scallops; additional designs within the parts of the drawing; excrescences, such as additional branches; and external additions, such as scribbling and doodling, writing, and other objects or designs unattached to the flower. A few abnormal *S*s added 'eyes' within the sepals or petals, as illustrated in Fig. 14. Quantitatively, additions are found in approximately twice as many abnormal as normal drawings.

In the analysis of *size errors*, those drawings which were markedly larger or smaller than the sample were tabulated. Height of stem, together with width of leaf or petals (or both) were employed in this classification. The minimum and maximum dimensions employed in the classification involved an actual increase or decrease of from 23 percent to 47 percent from the sample dimensions.[6] In terms of these standards, nine abnormal and two normal drawings were classified as larger, and 15 abnormal and six normal as smaller than the sample. More extreme examples, furthermore, are found among the abnormal, some of the smaller abnormal drawings, for instance, being so minute as to make the identification of parts difficult. An examination of *form errors* yielded no significant differentiation between abnormal and normal groups. Alterations of shape and of the relative proportion and position of parts of the drawing, lack of symmetry, and intellectual realism (indicated by transparency of overlapping parts) were all very common in both groups and the differences in frequency were slight and inconsistent. Such errors appear to be more closely related to the *S*'s skill and experience in drawing than to psychotic condition.

In regard to *coloring*, a number of clearly differentiating characteristics were found in the comparison of abnormal and normal drawings. Thus, complete refusal to color was obtained only from abnormal *S*s. Partly uncolored drawings were submitted by 10 abnormal and 2 normal *S*s, in the two latter cases the uncolored areas being small and apparently resulting from an oversight. Monochromes, colored entirely with a single crayon selected at random, were produced only by abnormal *S*s. Coloring only the outline in

[6] The same percentages obviously could not be used throughout because of the influence of the space available on the sheet.

part or all of the drawing was also observed only among abnormal *Ss*. 'Color area break-up,' or the use of more than one color in an area containing only one color in the model, was four times as common in the abnormal as in the normal group, and the peculiar application of color was approximately six times as common. The latter includes a variety of individual techniques in the application of crayons, such as scrawls, lines, dashes, check marks, or streaks. Some *Ss* applied the crayon with differential pressure within an area which was homogeneous in the sample. Others left large unfilled areas in the drawing, having made a few hasty scrawls in each section; a relatively mild example of this procedure is illustrated by Fig. 2. The use of crayons other than those selected by the experimenter for drawing No. 4 was observed in seven abnormal and two normal cases. In the two latter cases, however, black was used instead of grey, obviously as a result of careless observation. Other coloring errors, such as coloring loops or outlines in different colors, drawing directly in crayon rather than in pencil, filling in an area in pencil and then going over it in crayon, and several confusions in the placement of the given colors, although occurring somewhat more frequently in the abnormal group, did not yield very large or consistent abnormal-normal differences.

5. *Analysis of time data.*

It will be recalled (cf. Procedure) that both 'delay' and 'total time' were recorded for each of the four drawings, the former being the time elapsing from the end of the instructions to the *S*'s first marks on drawing or scrap paper, and the latter including the delay plus the time required to complete the drawing. In Table V will be found the mean, SD, and significance ratio of abnormal-normal mean differences for all four drawings, computed separately for male and female abnormal and male and female normal sub-groups. All time scores are reported in seconds.

In all four drawings, the mean *delay* is consistently longer in the abnormal than in the normal groups of both sexes, although the differences are not highly reliable, the significance ratios ranging from 0.87 to 2.75. Variability in the duration of delay is large throughout, but consistently greater in the abnormal groups. This finding suggests the presence in the abnormal groups of more impulsive performance with a minimum of planning at one extreme, as well as more resistant and irrelevant behavior at the other extreme.

Total time is likewise longer in the abnormal group in all comparisons, these differences being larger and more highly reliable than in the case of delay. The significance ratios of abnormal-normal mean differences in total time range from 2.52 to 6.78. Various factors

XUM

TABLE V

ANALYSIS OF TIME DATA *

Time in Seconds	Delay				Total Time			
	Abnormal Male	Normal Male	Abnormal Female	Normal Female	Abnormal Male	Normal Male	Abnormal Female	Normal Female
Drawing No. 1:								
Mean	41.47	34.72	48.48	33.67	691.28	382.47	693.78	447.02
SD	72.29	66.51	91.06	41.95	771.30	338.84	242.29	407.38
Significance ratio of abnormal-normal mean difference	0.87		1.92		4.82		6.78	
Drawing No. 2:								
Mean	66.18	50.76	89.77	51.62	498.23	313.46	590.73	386.96
SD	100.00	84.97	153.56	82.17	471.54	281.06	112.10	418.01
Significance ratio of abnormal-normal mean difference	1.50		2.75		4.27		6.12	
Drawing No. 3:								
Mean	29.39	17.39	30.56	17.35	503.09	320.33	493.17	378.82
SD	61.73	28.38	57.55	24.93	487.70	251.81	469.08	318.57
Significance ratio of abnormal-normal mean difference	2.23		2.62		4.21		2.52	
Drawing No. 4:								
Mean	22.94	14.17	16.34	12.45	957.09	743.69	937.05	782.89
SD	49.05	17.33	26.81	15.29	647.73	411.18	523.64	465.95
Significance ratio of abnormal-normal mean difference	2.13		1.59		3.51		2.80	

* In some of the measures reported in this table, *N* is less than 170, since a number of *S*s did not produce all four drawings.

may have contributed to the slower performance of the abnormal group. Continued resistance in the course of the drawing, tendencies to perseveration and over-elaboration, as well as the institutional situation itself, must be considered. The institutionalized *S* has more free time and less pressure from other activities. He also shows, in general, more concern regarding the 'test' which he is taking because of the bearing which he believes it may have upon the disposition of his case. Finally, variability in total time is again very large in all groups. In all but two comparisons, furthermore, variability is higher in the abnormal groups, indicating wider individual differences in this respect among the abnormal. The two exceptions are found in drawings No. 1 and No. 2, in which the abnormal women show a lower variability than normal women. Such a finding, coupled with the higher mean total time of the abnormal

women, indicates that the majority of abnormal women required an excessively long time to complete these drawings. With drawings No. 3 and No. 4, on the other hand, the more restricted task led to a more rapid performance among some of the abnormal women, thus raising the variability. The abnormal men showed greater variability than the normal men in all drawings.

6. *Survey of special peculiarities.*

In addition to the quantitative analyses of subject-matter, general technique, and time scores, which have been summarized above, each drawing was examined in reference to a check list of 'special peculiarities' assembled from the literature (cf. **1, 2, 3, 4**) as well as from the writers' own observations in their investigation of spontaneous drawings (**5**) and from a preliminary examination of the drawings obtained in the present study. For brevity, the term 'special peculiarities' will herein be employed to designate the characteristics in this check list.[7]

In general, the results of this analysis of each of the four drawings are mutually corroborative. Those characteristics yielding an appreciable abnormal-normal excess in one drawing, do so in all four drawings. Inconsistencies are found only when abnormal-normal differences are slight and negligible. Such agreement in the results obtained with the four drawings is all the more significant because each of the drawings was analyzed independently by the writers and no comparisons were made between the findings on each of the drawings until the entire tabulation was complete.

None of the items in the check list, on the other hand, occurs with a sufficiently high frequency in the abnormal group to be regarded as a regular feature of abnormal drawings. Each of the special peculiarities in the check list was observed in only a small minority of the abnormal sampling. When the frequencies of each special peculiarity in each of the four drawings by the 340 abnormal and 340 normal *S*s were tabulated, nearly 60 percent of such frequencies were found to be under 10.[8] Only 2 percent of the frequencies were over 100, and 4 percent were between 50 and 100. The larger frequencies, furthermore, were obtained in those characteristics which did not differentiate between abnormal and normal groups. It is apparent, therefore, that absence of any such characteristic from a drawing cannot be used as an indication of normality, even when the given characteristic yields a large, significant difference in frequency between abnormal and normal groups. Conversely, the presence of

[7] The check list contains 92 items and will be reproduced *in toto* in the forthcoming book to which reference has already been made.

[8] Exclusive of zero frequencies.

XUM

any one of these special peculiarities is not conclusive evidence of abnormality. Very few of the items in our check list were completely absent from the normal drawings. Those items having consistently zero frequencies in the normal group, furthermore, usually had very low frequencies (often two or three cases) in the abnormal group. Because of the rare occurrence of such characteristics, the zero frequencies in the normal group might be attributable to sampling error. Since, therefore, their interpretation is doubtful, no special consideration will be given to those characteristics appearing only in the abnormal drawings. Rather, *all* characteristics will be treated in terms of the size of the difference between abnormal and normal frequencies and in terms of the consistency of this difference in the four drawings.

A number of items in our check list yielded *negligible and inconsistent differences* [9] in frequency between the abnormal and normal groups. Among such items is to be found the use of scrap paper, which, although apparently related to the specificity of the task (its frequency being highest in drawing No. 1 and lowest in drawing No. 4), shows only slight and inconsistent differences between abnormal and normal groups. Voluntary change in the object drawn, as a result of an acknowledged inability to complete the first attempt, also shows no abnormal-normal differences. Going over the drawing repeatedly in pencil or crayon, thereby producing conspicuous dark masses and nearly wearing out the paper, likewise failed to differentiate between the two groups. The portrayal of allegories and myths was rare in either group, as were the 'mandala'-like designs of concentric circles reported by Jung (cf. **2**, pp. 13–14), in his discussion of abnormal drawings.

Many characteristics of technique proved to be more closely allied to lack of drawing skill and experience than to psychotic condition. Among them are to be listed various errors of perspective, spider-like fingers or toes, radiating lines around the sun, the use of lines to indicate the direction of motion of the objects portrayed, and the combination of full-face and profile views in a physically possible but unlikely posture. In the same category is found intellectual realism, or the portrayal of objects as they are known to be rather than as they appear in a single view, illustrated by transparency of overlapping parts, two eyes in a profile view, and the like. A clear example of intellectual realism is reproduced in Fig. 3, showing the simultaneous portrayal of front, back, and side views of

[9] No item in this category yielded a statistically reliable difference (critical ratio of 3.00 or more) between abnormal and normal percentages in any of the four drawings; inconsistencies in relative number of abnormal and normal subjects from one drawing to another were common and the differences in frequency were slight.

a house, as well as an internal view with table and chairs. Stylization and anthropomorphism, or the distortion of objects in the direction of greater resemblance to abstract, geometric forms (cf. Fig. 8) in the former, and to the human face or figure in the latter, occurred with very low, inconsistent frequencies in either group. An approximately equal number of abnormal and normal *Ss* perceived the branch of the stem in drawing No. 4 as the head of a snake, the resemblance being accentuated in a number of cases by the insertion of an eye. The use of a characteristic, idiosyncratic type of line, such as a broken, wavy, angular, scrawled, or sketchy line, showed slight, inconsistent abnormal-normal differences. In those drawings in which no sample was furnished (No. 1, No. 2, No. 3), micropsia, or the production of very small drawings, and its opposite, macropsia, also yielded negligible differences. An example of macropsia, obtained in drawing No. 2, is reproduced in Fig. 7. It is interesting to recall the more significant differences in reference to size obtained with drawing No. 4, in which a sample design furnished a guide for the size of the drawing. Under those conditions, marked size discrepancies were definitely more common and more extreme among the abnormal.

In reference to coloring, refusal to color was fairly common in both groups, such refusal being frequently shown by *Ss* who felt the use of crayons to be somewhat childish or who believed their efforts were too poor to warrant coloring. Making an outline drawing in color, filling in an area in pencil and then going over it in crayon, and the use of black crayon over pencil all yielded negligible abnormal-normal differences. Small, inconsistent differences were likewise obtained in the sparing use of color and in the application of color to a part of a drawing in such a way as to be incompatible with the entire pattern, the subject responding to isolated lines without reference to their representative value in the whole object.

We may consider in a second category those items yielding a *moderate but consistent excess in abnormal frequency*,[10] placing into this category also those characteristics occurring in a small number of

[10] Although the abnormal-normal differences were more conspicuous in the items now to be discussed than in those in the preceding category, they were not sufficiently large, in most cases, to be statistically significant. Only five out of a total of 109 abnormal-normal comparisons in the present group yielded critical ratios of 3.00 or more. Cognizance must also be taken of the fact that the occurrence of many of these characteristics was artificially limited or precluded in one or more of the four drawings by the nature of the task, as assigned in the instructions. This explains the inclusion of certain characteristics with fairly low frequencies in the present category. No characteristic in the present category showed reversals in the relative abnormal-normal frequency from one drawing to another, with the exception of a slight discrepancy in 'excessive productivity' in drawing No. 3. Since, however, the instructions in all but the first drawing were such as practically to preclude the production of several drawings, such an inconsistency in drawing No. 3 is probably insignificant.

XUM

abnormal drawings and completely absent in the normal. Among the latter characteristics may be mentioned: the discrete application of color to each element in a homogeneous area, such as bricks in a wall, although all are colored alike; obsessive concern with numbers, as in repeated counting of, for example, rungs in a ladder and insistence upon the inclusion of a specific number of them; the production of paranoid maps or plans of 'inventions;' the portrayal of cruelty and suffering; the repeated use of a stereotyped individual element or 'trade mark;' and submitting a single line or scrawl when directed to draw.

Excessive resistance was manifested by a consistently larger number of abnormal than normal *S*s in each of the four drawings, the difference in frequency increasing as the task became more specific. Thus the abnormal-normal difference in frequency was greatest in drawing No. 4. Deliberate destruction of the drawing by scrawling over, tearing, or crumbling occurred more often among abnormals. Excessive productivity, or the production of several drawings when only one was requested, on the other hand, was also more common among the abnormal. Very light as well as very hard pressure on the crayons or pencil occurred with consistently greater abnormal frequency. Turning the sheet from vertical to horizontal position and placing the drawing in an unusual position on the page, also more common among the abnormal, are both suggestive of more individualistic and less compliant behavior. The more frequent portrayal by the abnormal of objects visible in the immediate environment, on the other hand, may indicate a response similar to the 'stimulus-bound' drawings described by Becker (cf. 4, p. 192), in which the *S* elaborates a design around a printed letter or other mark accidentally present on the paper. In both cases, the *S*'s drawing behavior is influenced by stimuli fortuitously present in the immediate surroundings. The portrayal of sexual and of religious themes, though infrequent in either group, shows a slight excess among the abnormal.

The tendency to make crude, schematic, simplified drawings was consistently more common among the abnormal. The greater fullness of realistic detail in the normal drawings is further indicated by the more frequent portrayal of clothing, beard, motion or special pose, and accessory objects in the drawings of a man by normal subjects. Exaggerated meticulousness of detail, on the other hand, was observed with greater frequency in the abnormal drawings. Excessive reliance on perceptual filling, with omissions and blank spaces which the observer can fill in perceptually (cf. Fig. 9), as well as 'scatter' or inconsistency in fullness of detail portrayed in different parts of the drawing (cf. Fig. 12), was also more common among the abnormal.

With reference to coloring, unnatural coloring of objects was somewhat more frequent in the abnormal drawings, although not absent among the normal. Also more common in the abnormal group was the practice of drawing directly in crayon and ignoring the suggestion that pencil be used, a fact which may corroborate the greater tendency to resistant and individual behavior previously noted. Other coloring peculiarities which were somewhat more common among the abnormal are the use of an excessive amount or variety of color, coloring only the outline of a drawing previously made in pencil, and the use of white crayon over the white paper.

Finally, a number of peculiarities of technique occurring in relatively few abnormal or normal drawings, but with consistently greater frequencies among the former, include: lack of symmetry; windows drawn flush with the edge of the house; 'tadpole' men, a crude, disproportionate, bodiless drawing of a man in which the legs are attached directly to the head (cf. Fig. 12); overlapping function of lines, the same line representing parts of different objects, such as arms of one figure and legs of another; overlapping figures in space, the same area in the drawing serving a multiple function in the representation of more than one object; the portrayal of 'frames' around a drawing; and '*horror vacui*,' a term coined by Prinzhorn (cf. **3**, p. 14, 61) to describe the *S*'s tendency to fill the entire sheet with his drawing, leaving no blank spaces. An illustration of '*horror vacui*' in drawing No. 2 is given in Fig. 7.

In the third and last category we may include all of those items in the check list in which a *consistently large difference* [11] in frequency was observed between abnormal and normal groups, as well as those items with zero frequencies in the normal group and a sufficiently large frequency in the abnormal to justify serious consideration. These are the drawing characteristics most highly diagnostic of abnormality, although it should be remembered that none of the frequencies was in itself large enough to represent more than a small minority of abnormal cases. Among the characteristics in the present category is to be found 'blotting-over,' or repeatedly going over the drawing in pencil or crayon with the resulting obliteration of original forms or colors. The drawing must usually be 'rescued' by the experimenter in such cases in order to prevent its complete obliteration. A relatively mild form of 'blotting-over' is illustrated by the flower pot and leaves in the drawing reproduced in Fig. 1. Failure to follow

[11] No item in the present category showed a reversal of relative abnormal-normal frequency from one to another of the four drawings. Only characteristics occurring in more than five abnormal *S*s in at least one drawing were included. Even in the present group of characteristics, however, only about one third of all the abnormal-normal comparisons yielded critical ratios of 3.00 or more; in many of the other comparisons, the frequencies are too small to yield a statistically reliable difference.

directions, also observed with much greater frequency in the abnormal group, was manifested in various ways, such as deliberately producing a drawing not representing danger in drawing No. 2, portraying an object other than a person in drawing No. 3, or drawing a woman when directed to draw a man.[12] The portrayal of delusional and hallucinatory ideas, as well as drawings showing free association and flight of ideas, occurred exclusively among the abnormal and with fairly high frequencies. The exaggerated use of symbolism, especially symbolism of a highly individual nature, was also much more common in the abnormal group.

Stereotypy of content or technique in different drawings, incoherence or chaotic organization as well as disproportion or displacement of parts of a drawing, the inclusion of scribbling or scrawling lines, omission of essential details, and decorative overelaboration, all show consistently large abnormal-normal differences in frequency. An example of incoherence, i.e., the inclusion of irrelevant items, is given in Fig. 1, in which musical notes and circles containing crosses have been added haphazardly to a drawing of a plant. Especially large is the difference in frequency of perseveration, illustrated by 'cauliflower' and 'grape-like' formations and excrescences, the repetition of any identical detail or object within the same drawing (cf. Fig. 2), or the continued production of tiny lines in pencil or crayon with little or no discernible effect on the drawing. In extreme cases, the *S* will continue the latter procedure indefinitely until the paper is removed.

Among the most interesting specimens of abnormal drawings—and those which have received greatest attention in the literature—are those showing gross distortion, the objects portrayed being scarcely recognizable. Although by no means absent from the normal group, especially among *S*s of low educational level and little experience in the use of paper and pencil, such drawings occur with much greater frequency among the abnormal. Examples of such drawings, taken from the abnormal group, are reproduced in Figs. 4, 5, 8, 9, 10, 11, 15, and 16. Fig. 4, a 'free choice' drawing, is described by the subject as "a Pekingese dog." Fig. 5 portrays "a man drinking," produced by the *S* as a representation of danger (drawing No. 2); the man is colored entirely in black crayon. Figs. 8 to 11, inclusive, were executed in response to the directions to draw a man (drawing No. 3). Attention is called especially to the over-simplification of the form in Fig. 8, the 'open face' in Fig. 9, and the amorphous and wraith-like forms in Figs. 10 and 11, respectively. Figs. 15 and 16 show idiosyncratic distortion and over-simplification of the sample design to be copied in drawing No. 4.

[12] Drawings of a child, rather than of an adult, however, were made by no abnormal and only two normal *S*s.

A fairly heterogeneous group of drawings which can be broadly described as indicating 'unusual perception' of objects, is also much more common among the abnormal. In a few cases, actual perceptual disorders are involved, as in the characteristic drawings produced by a patient with traumatic psychosis; these drawings show marked dislocations and reversals of parts of the face, the representation of a face within a face, 'chain drawings' in which the body of one figure becomes the face of a larger one, and similar disruptions. Bizarre conceptions of the act of drawing, as well as the confusion of different levels of abstraction, are illustrated in other cases. Thus, one *S* drew head, trunk, and legs, but wrote the word "feet" at the proper place instead of drawing the feet. Another wrote the word "Man" as soon as he received the instructions for drawing No. 3. When told to show how the man looked, he added the word "Place," saying, "Now I've done what you asked me—*here's* the man." In response to the directions to draw a man, one patient made two crosses with an ellipse under each, explaining that she had "drawn two men in Chinese." When asked to "draw the man in English," she was able to produce a creditable profile. A patient who had made a copy of the box of crayons for his 'free choice' drawing, began to color the box rather than his drawing. In a drawing of danger, another patient made two 'open' heads, the features being drawn without an enclosing outline and the bodies being difficult to distinguish from the heads. The *S* explained that these figures represented two prostitutes, but since he wanted to draw twelve, he added the number "10" to his drawing.

Large abnormal-normal differences are found in the drawings showing an inconsistency in amount or distribution of color in different parts, as well as the peculiar, idiosyncratic application of color in streaks, splotches, lines, dashes, and the like. Especially conspicuous are the abnormal-normal differences with reference to the inclusion of writing in the drawings. A few abnormal *S*s produced only writing when directed to draw; a much larger number added writing to their drawing; and a considerable number appended elaborated signatures with geometric or decorative designs, colors, titles, names of places, etc. Certain qualitative differences also appear in the written matter added by abnormal and by normal *S*s. Thus among the normal *S*s, the writing usually consisted of the title of the picture, or labels on parts, whereas the abnormal frequently inserted irrelevant words or phrases, numbers, letters, or long, rambling passages.[13]

(Manuscript received January 20, 1944)

[13] Neologisms, or the coining of new words, were found in only four cases, all abnormal.

References

1. Anastasi, A., & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal: I. Historical and theoretical background. *J. gen. Psychol.*, 1941, **25**, 111–142.
2. Anastasi, A., & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal: II. Approaches and interrelationships. *Ann. N. Y. Acad. Sci.*, 1941, **42**. Pp. 106.
3. Anastasi, A., & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal: III. Spontaneous productions. *Psychol. Monog.*, 1940, **52**, No. 6. Pp. 71.
4. Anastasi, A., & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal: IV. Experimental investigations. *J. gen. Psychol.*, 1941, **25**, 187–237.
5. Anastasi, A., & Foley, J. P., Jr. An analysis of spontaneous artistic productions by the abnormal. *J. gen. Psychol.*, 1943, **28**, 297–313.

XUM

AN EXPERIMENTAL STUDY OF THE ROLE OF THE EGO IN WORK. II. THE SIGNIFICANCE OF TASK-ORIENTATION IN WORK

BY HELEN BLOCK LEWIS AND MURIEL FRANKLIN *

In a previous paper (**5**), an experiment was reported in which it was demonstrated that the resolution of tension-systems in work can occur as a result of another person' activity as well as one's own. Recall of interrupted (partner-completed) and self-completed tasks performed in a coöperative work relationship was equally frequent. In Zeigarnik's experiment (**10**), which was the prototype for our Coöperative Work Experiment (CW), interrupted tasks were recalled about 60 percent more often than completed ones.

Four additional experiments, ramifying from the results of the CW Experiment, will be reported in this paper. For a detailed description of the tasks used, and of general procedure, the reader is referred to the preceding paper (**5**) in this series.

Experiments I and IA

In the Coöperative Work Experiment, a ratio of 0.88 was obtained between average recall of interrupted (partner-completed) and self-completed tasks. This ratio was in striking contrast to Zeigarnik's (corrected) ratio of 1.61 (see Marrow, **6**). The possibility arose that the ratio of 0.88 was a function of something unknown about our experimental conditions at Brooklyn College. It became necessary to repeat Zeigarnik's experiment in our laboratory, using 18 tasks nearly identical with those employed in the CW experiment.[1] The repetition of Zeigarnik's experiment in Experiment I was designed, then, not so much to check on her results (which have been many times substantiated), but rather on our own tasks and conditions.

A. *Procedure*

In Experiments I and IA, each *S* worked alone, to perform the same 18 tasks described in the preceding paper. Nine tasks were interrupted by *E*'s saying: "I'll take that now," or "We'll do the next one now"; nine tasks were completed by the *S* without interruption. As previously, of course, the tasks were rotated between the interrupted and completed conditions.

It quickly became apparent that the attitude of the *S* toward the experiment was of prime importance in recall. Two sets of preliminary instructions were therefore developed in an at-

* These experiments were performed in the Brooklyn College Psychology Laboratories and were supported by Grants-in-Aid from the Social Science Research Council. The assistance of the Council is acknowledged with gratitude. The senior author is responsible for the text of this paper.

[1] Perfect identity was not always possible. Some tasks which took two people five minutes to do had to be cut in half so that one person would not need more than five minutes for completion. A smaller jig-saw was used, for example, half the number of sets of papers for stapling, etc. Only one task had its character changed—the Limerick (see **5**, p. 120).

tempt to govern the *S*'s attitude. In Experiment I, an intense effort was made to disabuse the *S* of any notion that a 'test' of him was being performed. The instructions ran:

> "I am planning some experiments for next semester, and would much appreciate your help in finding out something about the materials I want to use. You are a kind of preliminary 'guinea pig' who will tell me something about these materials. Just do the tasks I have prepared so that I can find out about them. You see, of course, that this isn't at all a test of *you*. It's a test of the tasks. You are in no sense on the spot."

In Experiment IA, these instructions were considerably less pointed:

> "I have some tasks here which I should like to have you do. Please work any way you like. This is in preparation for some experiments I want to perform next semester. This is a kind of preliminary."

B. *Results*

The differences in attitude apparent in the first few *S*s were accentuated by these instructions. *S*s in Experiment I were more at ease, talked during the work, expressed keen interest in some of the tasks and were disappointed at interruption. They were task-oriented. *S*s in Experiment IA, on the other hand, were obviously ill at ease and much more silent except for occasional outbursts of self-derogatory comments at the time of interruption. "How did I do?" was their most frequent and insistent question during the interview. They reported feeling 'self-conscious,' wondering what the *E* was really testing and felt keenly their inability to complete such 'simple' tasks. "If it wasn't a test of intelligence, then surely [the *E*] must be testing some important aspect of [their] personality." "Please tell me my results and tell me I'm not a complete moron," said one subject. "Every time you took it away," said another subject, "I felt: 'why couldn't I finish'? Because I'm such a poke." Another subject reported feeling a great deal of 'pride in completion,' because that meant he had done well 'on the test.' These subjects were not task-, but ego-oriented.

The recall ratios in Tables I and II show these differences clearly. It will be seen at once that the subjects in Experiment I recall more interrupted tasks, while the subjects in Experiment IA recall more completed than interrupted tasks. The ratio between the average number of interrupted tasks and the average number of completed tasks recalled in Experiment I is 1.74, which compares favorably with Zeigarnik's (corrected) ratio of 1.61 and with Marrow's obtained ratio of 1.57. If we assume that 1.60 is the 'true' ratio which occurs in recall when task-completion tension systems are interrupted (this is, in fact, Marrow's median ratio), then we may expect that, if this factor is operating in our results, our distribution around 1.60 should not be different from Marrow's. In 30 cases Marrow obtained 19 ratios of 1.60 or above, and 11 ratios below. We obtained, in 12 cases, 8 ratios of 1.60 or above and 4 ratios below. Applying

TABLE I

SHOWING RI, RC, RT, RI/RT, RI/RC FOR EACH *S* IN EXPERIMENT I

Subj.	RI	RC	RT	RI/RT	RI/RC
1	8	2	10	.80	4.00
2	6	2	8	.75	3.00
3	8	4	12	.75	2.00
4	4	2	6	.67	2.00
5	6	3	9	.67	2.00
6	6	3	9	.67	2.00
7	7	4	11	.64	1.75
8	8	5	13	.62	1.60
9	6	4	10	.60	1.50
10	8	6	14	.57	1.33
11	4	4	8	.50	1.00
12	4	4	8	.50	1.00
Means	6.25	3.58	9.83	.64	1.93

RI = number of interrupted tasks recalled.
RC = number of completed tasks recalled.
RT = total number of tasks recalled.

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RC}} = 1.75.$$

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RT}} = 0.64.$$

TABLE II

SHOWING RI, RC, RT, RI/RT, RI/RC FOR EACH *S* IN EXPERIMENT IA

Subj.	RI	RC	RT	RI/RT	RI/RC
1	4	4	8	.50	1.00
2	4	5	9	.44	.80
3	4	5	9	.44	.80
4	4	5	9	.44	.80
5	3	4	7	.43	.75
6	5	7	12	.42	.71
7	5	8	13	.38	.63
8	4	7	11	.36	.57
9	4	7	11	.36	.57
10	4	9	13	.31	.44
11	2	5	7	.29	.40
12	2	6	8	.25	.25
Means	3.75	6.00	9.75	.39	.64

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RC}} = 0.625.$$

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RT}} = 0.38.$$

the chi-square test to this difference, we find a chi-square of 0.0415, which, with one degree of freedom, yields a *P*-value between 0.80 and 0.90. Our distribution does not differ significantly from Marrow's.

The recall results in Experiment IA are, on the other hand, radically different from Zeigarnik's, Marrow's and the results of Experiment I. Here the ratio between the average number of interrupted and the average number of completed tasks recalled is

0.625. Applying the chi-square test to the difference between the distribution of ratios in Experiment IA and Marrow's distribution yields a *P*-value of <0.01, indicating that the difference is significant. The difference between Experiments I and IA also yields a *P*-value <0.01, indicating a significant difference.

Analysis of the recall ratios for individual tasks in Experiment I and IA shows clearly that the results of both experiments are not a function of the particular tasks used. Table III shows that in Ex-

TABLE III

SHOWING TASK-ANALYSIS OF RECALL IN EXPERIMENT I

No.*	Task	C	RC/C	I	RI/I	RT/T
3	Clay House	5	.80	7	1.00	.92
12	Jig-saw	6	.83	6	1.00	.92
8	Packing	6	.67	6	1.00	.83
1	Winding	6	.67	6	.83	.75
15	Rearranged Sentences	3	.67	9	.78	.75
2	Anagrams	6	.50	6	.83	.67
5	Cutting and Pasting	7	.43	5	1.00	.67
7	Map	6	.33	6	.83	.58
10	Stick Problem	6	.50	6	.67	.58
18	Letter	4	1.00	8	.38	.58
6	Adding	6	.50	6	.50	.50
4	Limerick	6	0.00	6	1.00	.50
9	Alphabetizing	6	.33	6	.67	.50
16	Braiding	6	.50	6	.33	.42
17	List	5	.20	7	.57	.42
13	Vowels	6	0.00	6	.33	.17
14	Stapling	9	0.00	3	.33	.08
11	College Plan	7	0.00	5	0.00	0.00

* The tasks have been arranged in order of recall value. The numbers are simply the arbitrary number assigned to the tasks. They do not represent serial order of presentation.

C = Number times presented as completed tasks.

$$RC/C = \frac{\text{Number completed tasks recalled}}{\text{Number completed tasks presented}}.$$

I = Number times presented as interrupted task.

$$RI/I = \frac{\text{Number interrupted tasks recalled}}{\text{Number interrupted tasks presented}}.$$

$$RT/T = \frac{\text{Total number tasks recalled}}{\text{Total number tasks presented}}, \text{ i.e., recall value of task.}$$

periment I, the majority of tasks were recalled more often in the interrupted than in the completed condition. Table IV shows that in Experiment IA, the identical tasks were recalled more often in the completed than in the interrupted condition.

One may postulate that for the *S*s in Experiment I tension systems were aroused to complete the tasks. Interruption left unresolved some of these task-completion tension systems. For the *S*s

TABLE IV

SHOWING TASK-ANALYSIS OF RECALL IN EXPERIMENT IA

No.*	Task	C	RC/C	I	RI/I	RT/T
3	Clay House	7	.72	5	1.00	.83
8	Packing	7	1.00	5	.60	.83
2	Anagrams	5	1.00	7	.57	.75
15	Rearranged Sentences	4	.75	8	.75	.75
18	Letter	8	.63	4	1.00	.75
12	Jig-saw	6	.67	6	.67	.67
16	Braiding	6	1.00	6	.33	.67
1	Winding	6	.83	6	.33	.58
9	Alphabetizing	6	.83	6	.17	.50
10	Stick Problem	4	.50	8	.50	.50
4	Limerick	6	.50	6	.33	.42
7	Map	7	.57	5	.20	.42
13	Vowels	6	.83	6	0.00	.42
14	Stapling	7	.57	5	.20	.42
17	List	6	.50	6	.33	.42
6	Adding	5	.20	7	.43	.33
11	College Plan	9	.33	3	.33	.33
5	Cutting and Pasting	6	.33	6	0.00	.17

* See Footnote to Table III.

in Experiment IA, on the other hand, tension systems were aroused to do well, to protect their ego-status by performing the tasks. Interruption meant, primarily, a loss of ego-status, and only secondarily a suspension of the tension-system for task-completion.

The results of these two 'control' experiments show, then, that interruption does result in recall advantage for interrupted tasks when the tension-systems aroused by the entire experimental situation are predominantly task-completion tension-systems and not ego-enhancement systems.

The relatively great frequency of ego-enhancement cases in Experiment IA was not only a function of our instructions, since some of these cases appear in Experiment I (*Ss* No. 11 and 12). It was undoubtedly a function of the strong system of reward for personal achievement in which Brooklyn College students have for the most part been bred. It was for this reason that special instructions were necessary to keep our *Ss* focussed on the tasks rather than on their own IQ's.

Zeigarnik reports finding some occasions on which the *S* felt he could not complete a task and that interruption signified his 'inferiority' or failure. All Zeigarnik tells us about these occasions is that out of 40 such tasks, only 32 percent were recalled, a much lower percentage than the usual 60 percent recall of interrupted tasks. Zeigarnik believes that a 'quasi-need' to complete such a task did exist, so that it would be recalled if it were not subject to 'repression.' "One cannot escape the conclusion," she writes, "that

a special encapsulation, a repression, . . . made recall more difficult." (**10**, p. 77, present authors' translation.)

Abel's (**1**) study of the relation of the Schneider index to recall of interrupted and completed tasks also revealed that *S*s with low Schneider index (usually associated with neurosis) recalled completed rather than interrupted tasks. She suggests that these *S*s were ego- rather than task-oriented, and hints that repression may be at work to prevent recall of interrupted (failure) tasks. Rosenzweig, in a series of papers (**7**, **8**, **9**), especially his most recent report, has also advanced experimental evidence in support of the thesis that failure experiences are repressed in recall when the person is 'defending his ego.'

Special Forces of Repression?

Experiment I showed that we could obtain in our laboratory results similar to those obtained by Zeigarnik. These results are radically different from the results of the CW experiment. The chi-square test applied to the difference between distributions yields a *P*-value of <0.01. Now the question arises whether the factors operating to create a ratio of 0.88 in the CW Experiment were not the same factors operating in Experiment IA to create a ratio of 0.625. The median ratio in Experiment IA is 0.67. If we assume that 0.67 is the 'true' ratio obtained when ego-enhancement tension-systems have been disturbed, then if this factor were operating in the results of the CW Experiment, we should expect similar distributions of ratios in both Experiment CW and Experiment IA. The chi-square test applied to the difference between the two distributions yields a *P*-value of 0.05, which indicates that there are only five chances in 100 that the difference between the two experiments is not significant.

The protocols of the *S*s make perfectly clear the real difference between the two experiments. In the CW Experiment, most of the *S*s were simply not concerned with questions of their own personal success or failure in performing the tasks. In Experiment IA, this was the most pressing question for the majority of *S*s. Most of the *S*s in Experiment IA were behaving, in other words, like *S* No. 14 in the CW Experiment, and presumably like Zeigarnik's 'special forces' *S*s. This leads directly to the following series of hypotheses.

1. Whenever the person is using tasks as a means of ego-enhancement, whenever the relation between the task and the goal is *P*—— task——→ ego-status, then interruption of the tasks is more likely to be regarded as a blow at ego-status. Since the goal is enhanced ego-status, and interruption prevents fulfillment of that goal, interruption is likely to give rise to feelings of failure.

2. In such cases, the goal is not task-completion, which is, rather, a means to the goal; therefore interruption should not result in a difference in magnitude between the task-completion tension-systems for interrupted and completed tasks. In Lewin's terms, (4) $ts_c = ts_i$, when ts_c represents tension-systems to recall completed tasks and ts_i represents tension-systems to recall interrupted tasks. The recall of interrupted tasks and of completed tasks should therefore occur equally, by chance, for ego-oriented *S*s. That is, more precisely, if the average total number of tasks recalled is about nine, or 50 percent of all tasks presented, then by chance interrupted tasks should contribute half to this total and completed tasks should contribute half to this total, for ego-oriented *S*s. In our Experiment IA we should expect a ratio of 1.00, resulting from about 4.5 interrupted and 4.5 completed tasks being recalled, and depending upon the equivalence of the task-completion tension-systems in the interrupted and completed conditions.

3. (*a*) But interruption is more likely to result in feelings of failure than is completion. (*b*) Conversely, completion is more likely to result in feelings of success, of ego-enhancement, than is interruption. Two possible explanations for the obtained ratio of 0.625 in Experiment IA are thus possible from this point on. The Freudian view and presumably Zeigarnik's "special forces of repression" would postulate that ego-wounding failure experiences would be avoided (repressed) resulting in the non-appearance of failure (interrupted) tasks in recall. In fact, one minimum condition set down by Freud (3) as required for the appearance of repression does exist in Experiment IA—the situation affords a threat to self-esteem. Obviously, the converse of this view would hold that the difference in recall arises not because ego-wounding experiences do *not* appear in recall, but because ego-satisfying experiences *do*.

Now, if ego-wounding experiences are being avoided, or pushed aside (repressed) in recall, then one should expect very few of the failure tasks to appear in the recall ratio. In other words, if repression is occurring in Experiment IA, then one should expect an 'abnormally' low percentage of interrupted (failure) tasks in recall. Such an abnormally low percentage of interrupted tasks is at least a minimum *statistical* criterion for the existence of repression.

If, on the other hand, repression is not at work, then one should expect a percentage of interrupted tasks which is only 'relatively' low— that is, relative to a higher percentage of completed tasks in recall. The recall ratio should, in other words, show a preponderance of ego-satisfying, i.e., completed tasks, without a serious absence of interrupted tasks.

The question immediately arises: what shall be considered evidence of 'abnormally' low percentage of interrupted tasks, of 'serious absence' of interrupted tasks in recall? We propose to use the pattern of recall in Experiment I as a statistical standard by which to judge the recall pattern in Experiment IA.

In Experiment I, 54.6 percent, or roughly speaking, 50 percent of all tasks presented were recalled. This is a characteristic finding in the Zeigarnik experiment. Of the actual total of 118 tasks recalled, 75 or 64 percent were interrupted tasks, while 43 or 36 percent were completed tasks. These figures agree well with both Zeigarnik and Marrow's independent findings. Marrow (**6**) found that 61 percent of all recalled tasks were interrupted and 39 percent completed. Zeigarnik's (**10**) results show 62 percent of interrupted and 38 percent of completed tasks recalled. Roughly speaking, then, interrupted and completed tasks stand to each other in the ratio of 6 : 4. The appearance of approximately 40 percent of completed tasks in recall can be regarded as a 'normal' or 'usual' occurrence in the Zeigarnik experiment.

In our Experiment IA, the ratio of incompleted to completed tasks recalled is exactly the reverse of Experiment I. Out of a total of 117 tasks recalled in Experiment IA, 72, or 62 percent, were completed tasks, and 45, or 38 per cent were interrupted tasks. The statistical magnitudes are obviously the same, only their positions are reversed. If we are justified in regarding roughly 40 percent of completed tasks as a 'normal' percentage in recall, then by the same token, 40 percent of interrupted tasks can be regarded as equally 'normal.' Judged by this standard, the recall of interrupted tasks in Experiment IA is not abnormally low.

Our data, then, do not offer any direct evidence for the existence of any "special forces of repression" operating on failure tasks in Experiment IA. They cannot, of course, be regarded as conclusive evidence against the existence of repression. It is theoretically quite possible that quite different processes at work in Experiments I and IA could still yield similar statistical magnitudes. But the data do suggest that extreme caution is necessary before the concept of repression is invoked to explain what is essentially a greater recall of success tasks rather than an actual absence of failure tasks in recall.

Our interpretation of the results of Experiment IA favors an hypothesis which suggests that the greater recall of completed tasks reflects a tendency for ego-enhancement experiences to appear in the recall of ego-oriented *S*s. We are unable to discover any special forces of repression in these *S*s. Rather, the *S*s appear to be geared to success, and so to recall more successful (completed) tasks.

XUM

This raises a question about 'pride in completion' which can be only very briefly treated here. It would be a mistake to assume that taking personal pride in task completion is *always* a reflection of ego-orientation. Ego-orientation is likely to bring with it pride in personal success, but pride in personal success can also occur without ego-orientation. A person who has just solved a difficult problem will feel pride in his success, although he did not attack the problem to win this 'pride.' What is important in this case is that so much of the person was involved in his struggle with the difficult problem—his intelligence, his emotions, his energy—the problem, in other words, was so central to the person, that solution results in satisfaction with the competent self. This self was, however, not necessarily focal in the person's activity—rather, the problem was focal.

Experiment II

It was possible, also, that the ratio of 0.88 obtained in the CW Experiment was simply a function of the failure of specific tension-systems to arise at all in our coöperative work. Zeigarnik found that when *S*s performed the tasks on a tour of the laboratory, just to see what sorts of things were done in a psychological laboratory, an average ratio of 1.03 was obtained. It might be argued that working jointly with another person prevents the assumption of responsibility for any particular task on the part of either worker, and particularly on the part of our *S*s in Experiment CW who felt slightly subordinate to the planted co-worker. It became necessary, therefore, to study the fate of tension-systems in coöperative work when both *S* and co-worker were interrupted and the tasks left incomplete.

A. *Procedure*

In Experiment II, the *S*s worked together with a planted co-worker on the identical 18 tasks used in Experiment CW. An attempt was made to conduct this experiment in as friendly and informal a work-atmosphere as had obtained in Experiment CW. This was not altogether successful, since the interruption was done by the *E* who was a teacher, rather than by the co-worker, a fellow-student, but it was, by and large, successful. The tasks were announced as in preparation for an experiment next semester, and the *S*s were obtained by asking them to help the planted co-worker "go through them." Both workers were interrupted by the *E*, who said: "I'll take that now." Half the tasks were thus left incomplete.

B. *Results*

Table V shows clearly that the recall ratios favor the recall of interrupted tasks. The ratio between the average number of interrupted and the average number of completed tasks is 1.50, as compared to 1.74 obtained in Experiment I. The *S*s in Experiment II obviously did not behave the way Zeigarnik's 'tourist' *S*s did. The difference between the distributions in Experiment I and Experi-

TABLE V

SHOWING RI, RC, RT, RI/RT, RI/RC FOR EACH *S* IN EXPERIMENT II

Subj.	RI	RC	RT	RI/RT	RI/RC
1	6	1	7	.86	6.00
2	7	2	9	.78	3.50
3	6	2	8	.75	3.00
4	5	2	7	.71	2.50
5	4	2	6	.67	2.00
6	5	3	8	.63	1.67
7	5	4	9	.56	1.20
8	5	4	9	.56	1.20
9	6	6	12	.50	1.00
10	2	4	6	.33	.50
11	3	6	9	.33	.50
Means	4.91	3.27	8.18	.60	2.10

$$\text{Ratio } \frac{\text{ave. RI}}{\text{ave. RC}} = 1.50.$$

$$\text{Ratio } \frac{\text{ave. RI}}{\text{ave. RT}} = 0.60.$$

ment II is not significant: $P > 0.75$. The task analysis presented in Table VI likewise shows a recall advantage for the interrupted condition in the majority of tasks.

TABLE VI

TASK-ANALYSIS—EXPERIMENT II

No.	Task	C	RC/C	I	RI/I	RT/T
2	Anagrams	4	1.00	7	.71	.82
3	Clay House	5	.40	6	1.00	.73
16	Braiding	7	.71	4	.75	.73
5	Cutting and Pasting	7	.71	4	.50	.64
14	Stapling	4	0.00	7	1.00	.64
17	List	7	.71	4	.50	.64
1	Winding	6	.17	5	1.00	.55
12	Jig-saw	5	.40	6	.67	.55
9	Alphabetizing	7	.71	4	.25	.55
15	Rearranged Sentences	4	.50	7	.56	.55
18	Letter	5	.20	6	.67	.45
6	Adding	8	.25	3	.67	.36
8	Packing	4	0.00	7	.56	.36
10	Stick Problem	5	.20	6	.33	.27
13	Vowels	5	.20	6	.17	.18
11	College Plan	6	0.00	5	.20	.09
7	Map	5	0.00	6	.17	.09
4	Limerick	5	0.00	6	0.00	0.00

The protocol of one *S* makes it clear that interruption operated in Experiment II in much the same way that it operated in Experiment I, or in the main body of Zeigarnik's experiment. This *S* said: "It made me sore when you interrupted me. It gave me such an unfinished feeling."

XUM

It is interesting that the ratio between average recalls should be slightly lower in this Experiment than in Experiment I. While the difference is not statistically significant, the pattern of difference in recall between the two experiments repays closer examination. In the first place, it is apparent that two cases of reversal of ratio, i.e., more completed than interrupted tasks, appear in Experiment II, while there are no such cases in Experiment I. These two *S*s were, as far as could be determined from their protocols, quite ego-oriented, behaving, in other words, like the *S*s in Experiment IA. Their presence in Experiment II would account for the lowered average recall ratio.

Now suppose these two *S*s are eliminated from the computation of results in Experiment II. The ratio between average recall of interrupted and average recall of completed tasks is now 1.88, a ratio even larger than that obtained in Experiment I.

One other difference between Experiment II and Experiment I is striking: fewer tasks are recalled in Experiment II than in Experiment I. The average number of tasks recalled in Experiment II is 8.18, while in Experiment I it was 9.83. In this respect, Experiment II resembles the CW Experiment in which only 7.92 tasks were recalled in all.

We may postulate that the tasks in Experiments II and CW were part of a total job—to help the co-worker. Since they were part of a total job, being done in a friendly and informal work-atmosphere, it is possible that a greater coalescence of tasks took place in Experiments II and CW than in Experiments I and IA. The boundaries between individual tasks were perhaps less rigid and the memory traces less distinct.[2] It thus became more difficult to recall the individual tasks and a lower total recall figure was obtained.

At any rate, then, coöperative work does not engender such an irresponsible attitude that no tension-systems to complete tasks are aroused. The recall ratio in the CW Experiment cannot be accounted for in this way. The crucial factor in the CW Experiment seems to have been that the interrupted tasks were completed by the coöperating worker and so regarded as finished.

Experiment III

This raises immediately the question of the role of task-completion in determining recall. If the concept of objective orientation has any meaning, if it is true that in some situations the goal of the individual is task-completion so that the completion of the task directly brings fulfillment of the goal, then the resolution of such task-com-

[2] Zeigarnik, it will be remembered, found that when the individual tasks became part of a larger whole, the ratios were reduced.

pletion tension-systems ought to be effected by *someone else's* completion of an interrupted tasks, *even in a non-coöperative situation.* Coöperative situations are simply situations in which it is easier for the objective orientation to occur. Or, put in another way, coöperative situations need as minimum requirement, objective orientation, but the effects of objective orientation ought to be clearly demonstrable in non-coöperative situations as well.

The problem raised here, of course, touches at the heart of motivation theory. The separateness of the person from the world around him is so obvious and axiomatic a fact about human beings that we naturally speak of a man's personal needs, wishes, or tension-systems, meaning by 'personal' just simply 'of the person.' But perhaps under some conditions, the needs of a person structure themselves to correspond with the structure of the objective situation, just as in perception what we see corresponds directly to the organization of the objective field. This is like a translation of Wertheimer's principle of isomorphism into the field of motivation. This interaction between personal needs and the external situation is quite direct, involving no intermediate step of referring the person's needs to the consideration of the person's ego.

Adler and Kounin (**2**) performed an experiment some years ago in which they attempted to show that the tension-system aroused by the assignment and assumption of a task is a *personal* tension-system. So they confronted a child with the alternative of resuming a task which he had begun and had not completed, or an identical task which he had not begun, but which was in an identical state of objective incompletion. The children chose to complete first the tasks which they had begun themselves.

Actually, this experiment of Adler and Kounin missed the crucial point. Their results do not demonstrate that a person's needs may not directly conform to the objective requirements of the situation, but only that when the objective requirements are *identical,* a greater need arises to complete a task which one has personally begun. I think there is a good chance that if Adler and Kounin had faced a child with an alternative of completing a task which he had begun and some different task requiring immediate action—such as righting a falling chair—the non-personal task might have won out.

Our experiment was designed to see what the effect of objective completion by another agent is upon recall of tasks.

A. *Procedure*

In this experiment, the person, working alone, performed the 18 tasks of Experiments I, IA, and II. Interruption was effected by *E*'s saying, "I'll take that now." *E* then moved the task from the *S*'s place at the work table and completed it while the *S* watched. Instructions for this experiment were otherwise identical with the instructions for Experiment I.

B. *Results*

Table VII shows the recall ratios for 23 *Ss*. It will be seen at once that the recall ratios are smaller than the ratios in Experiment I,

TABLE VII

SHOWING RI, RC, RT, RI/RT, RI/RC FOR EACH *S* IN EXPERIMENT III

Subj.	RI	RC	RT	RI/RT	RI/RC
1	7	3	10	.70	2.50
2	6	3	9	.67	2.00
3	7	4	11	.64	1.75
4	7	4	11	.64	1.75
5	5	3	8	.63	1.67
6	6	4	10	.60	1.50
7	7	5	12	.58	1.40
8	7	5	12	.58	1.40
9	4	3	7	.57	1.33
10	5	4	9	.56	1.25
11	5	4	9	.56	1.25
12	5	4	9	.56	1.25
13	6	5	11	.55	1.20
14	6	5	11	.55	1.20
15	6	5	11	.55	1.20
16	6	5	11	.55	1.20
17	7	6	13	.54	1.16
18	6	6	12	.50	1.00
19	4	4	8	.50	1.00
20	5	6	11	.45	.83
21	4	6	10	.40	.67
22	4	6	10	.40	.67
23	4	8	12	.33	.50
Means	5.60	4.70	10.30	.55	1.28

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RC}} = 1.20.$$

$$\text{Ratio}\,\frac{\text{ave. RI}}{\text{ave. RT}} = 0.54.$$

although the advantage in recall consistently favors the interrupted tasks. The ratio between the average number of interrupted tasks recalled and the average number of completed tasks recalled is 1.20. Table X shows a recall advantage for the interrupted condition in the majority of tasks.

The difference between the distribution of ratios in Experiment III and Experiment I is statistically significant: $P < 0.01$. The distribution of ratios is also significantly different from a distribution which might have been obtained if only chance factors were operating: P here is 0.02. The distribution is also most significantly different from the CW distribution of ratios: P is <0.04. In other words, the results in this experiment parallel the results of the CW Experiment more than the results of Experiment I.

Several possibilities can account for the fact that the ratio in Experiment III is slightly more than 1.00, but much lower than 1.74. The protocols of the *S*s make it clear that different types of forces were at play in different *S*s, so that 1.20 is a reflection of several different types of process.

For example, *S* No. 13 (ratio 6I/5C) seemed to feel that he was helping the *E* to finish the tasks. This *S* had actually to be prevented from physical participation in the *E*'s task completion. When he was thus physically prevented, he issued verbal instructions to the experimenter. In fact, when the *E* finished the clay house he said: "Oh, I was going to make a verandah, but our joint house is O.K." This *S* was annoyed by interruption and tried to 'race' the *E* to the point where she interrupted him. He had, he said, to "keep himself from being annoyed by interruption by remembering that it was part of the game." But once interruption did occur, he managed a vicarious completion by participating in the *E*'s work. Another *S* also makes this point. She said: "When you (the *E*) started to do them, I more or less did them with you and in a way completed them" (No. 14, 6/5).

Another *S* reveals a quite different process. She says: "I felt you were completing a task of your own. I did want to finish, but felt from the fact that it was an experiment that it was more important that you finish. I felt subordinate to whatever occurred" (No. 8, 7/5). Here the interruption is apparently not resolved by the *E*'s completion because at the point of interruption the task is no longer the *S*'s, but the *E*'s. The completion is not of one's own task, but of someone else's. More interrupted than completed tasks are therefore recalled.

For most *S*s, protocols show that interruption was an annoyance and that the *E*'s completion was unsatisfactory. The *S*s indicated that they would have preferred to finish themselves, but, as one *S* put it: "Your (i.e., the *E*'s) finishing was better than letting things hang in mid-air."

One group of three *S*s, finally, found the *E*'s completion not only unsatisfactory but a period of discomfort. For example, *S* No. 20 was so restless and ill-at-ease that he tried to leave the experiment on three separate occasions. He said: "I didn't feel natural when you interrupted me. Maybe it was because you saw I wasn't doing it right."

When the *S* is task-oriented, then, objective completion may offer some release of the task-completion tension-system. This release is effected either by actual vicarious experience, or because completion is "better than letting things hang in mid-air." In other cases, interruption signifies such a break in the task that it changes 'owner-

ship' and completion by another is not completion of my task, but of his. These two factors, together with three ego-oriented *Ss*, account for the obtained ratio of 1.20.

It is likely that objective completion is satisfactory for certain tasks and not for others. Task analysis for this Experiment (Table VIII) shows that there is some tendency for greater recall of self-

TABLE VIII

SHOWING TASK-ANALYSIS FOR EXPERIMENT III

No.	Task	C	RC/C	I	RI/I	RT/T
12	Jig-saw	10	.80	13	1.00	.92
2	Anagrams	13	.85	10	.90	.87
3	Clay House	13	.77	10	.90	.83
16	Braiding	12	.75	11	1.00	.83
8	Packing	11	.91	12	.75	.78
1	Winding	10	.90	13	.54	.70
15	Rearranged Sentences	6	.67	17	.65	.65
5	Cutting and Pasting	10	.40	13	.77	.61
6	Adding	11	.45	12	.75	.57
9	Alphabetizing	13	.54	10	.60	.57
18	Letter	13	.39	10	.70	.52
17	List	10	.50	13	.39	.44
4	Limerick	12	.33	11	.45	.39
7	Map	12	.42	11	.27	.35
10	Stick Problem	11	.27	12	.42	.35
13	Vowels	10	.30	13	.39	.35
11	College Plan	14	.29	9	.33	.30
14	Stapling	15	.20	8	.50	.30

completed tasks among the routine tasks; i.e., of five tasks which show greater recall of self-completed tasks, four are routine as against one non-routine. On the other hand, there is a greater difference in favor of recall of interrupted tasks among routine than among non-routine tasks. This may be purely a chance function of the greater number both of routine tasks and of tasks which were more often recalled in the interrupted condition. The relation of the potency of objective completion in effecting release of tension to type of task is a most important question for further experimentation.

It would seem reasonable to suppose that externally produced completion would be most difficult to produce satisfactorily with non-routine tasks, where the conclusion of the task was least rigidly defined by the nature of the task itself. In such a task, it might be almost impossible for the person who completed the task to hit upon exactly the same kind of solution which the task-performer had in mind. Next in order of difficulty would be routine tasks with an unstructured, arbitrary outcome, since external completion would not really matter, being only an additional number of lines added, or vowels crossed out, etc. Easiest tasks for the production of satis-

factory external completion would be the tasks where the outcome was both well-structured and clearly defined.

Experimental 'Controls'

Before final conclusions can be fully drawn there remain certain technical considerations about the conduct of the experiments which should be cleared away.

A. *Task Analysis*

As indicated in the discussion of each experiment, both in the preceding paper (**5**) and in this report, task analysis shows that the results obtained in each condition were not a function of the particular tasks used, but of the experimental conditions. In each condition the majority of tasks show the results of that condition. Thus, there is recall advantage for interruption in the majority of tasks in all conditions except Experiment IA and Experiment CW (see preceding report, Table II). In Experiment IA the same tasks show recall advantage under the completed conditions; in Experiment CW, the same tasks show recall advantage equally distributed between the interrupted and the completed conditions.

It was not always possible rigidly to adhere to a fixed plan of task presentation. If an *S* began to take an inordinately long time over a to-be-completed task it might have had to be converted on the spot to an interrupted one, and readjustment in plan made all along the line of tasks. Or a task being finished very quickly might have had to be converted into a completed task or else the *S* would have spent only a few seconds on it before interruption. This is a familiar mechanical difficulty with the Zeigarnik technique. By and large, however, tasks were presented as completed and interrupted an equal number of times.

B. *Serial Order of Presentation*

Four serial orders of presentation were developed:

Serial Order A	Tasks 1–18
Serial Order B	Tasks 18–1
Serial Order C	Tasks 9–1; 10–18
Serial Order D	Tasks 10–18; 9–1

Of these four serial orders, only two were carried through with sufficient frequency. C and D were not used often enough because the original number of subjects scheduled for each experiment was unexpectedly curtailed. For this reason, B and C have been combined to create serial order I, and A and D have been combined to create

serial order II. A and D have half the series in common; B and C have half the series in common.

Table IX shows the effects of serial order of presentation upon recall of tasks, combining all experiments except the CW Experiment.

TABLE IX

SHOWING EFFECT OF SERIAL ORDER OF TASK PRESENTATION ON RI, RC, RT FOR EXPERIMENTS I, IA, II, III COMBINED

Serial Order	N	Ave. RI	% RI (RI/I)	Ave. RC	% RC (RC/C)	Ave. RI/ Ave. RC	Ave. RT
A	21	4.90	.55	4.40	.49	1.11	9.30
B	23	5.69	.63	4.32	.48	1.32	10.01
C	7	4.70	.52	4.60	.51	1.02	9.30
D	7	5.14	.57	4.71	.52	1.09	9.85
	58						
Serial Order							
1. B–C	30	5.40	.60	4.40	.49	1.22	9.80
2. A–D	28	4.95	.55	4.46	.49	1.11	9.41

Ave. RI = average number interrupted tasks recalled.
Ave. RC = average number completed tasks recalled.
Ave. RT = average total number tasks recalled.

$$\%\,\mathrm{RI} = \frac{\text{total number interrupted tasks recalled}}{\text{total number interrupted tasks presented}}.$$

$$\%\,\mathrm{RC} = \frac{\text{total number completed tasks recalled}}{\text{total number completed tasks presented}}.$$

TABLE IXA

SHOWING EFFECT OF SERIAL ORDER OF TASK PRESENTATION WHEN EXPERIMENT IA IS OMITTED

Serial Order	N	Ave. RI	% RI (RI/I)	Ave. RC	% RC (RC/C)	Ave. RI/ Ave. RC	Ave. RT
A	13	5.70	.63	4.00	.44	1.43	9.70
B	22	5.71	.63	4.16	.46	1.37	9.87
C	5	4.80	.53	3.20	.35	1.50	8.00
D	6	5.67	.63	4.68	.52	1.21	10.35
	46						
Serial Order							
1. B–C	27	5.56	.62	4.00	.44	1.39	9.56
2. A–D	19	5.70	.63	4.20	.47	1.36	10.00

It will be seen that the difference between %RI recalled in series I and II is small (five percent) while the difference between %RC

recalled in Series I and II is zero. The difference between %RI recalled in the two serial orders is statistically unreliable—D/σ_D is 1.00, and is accounted for by the accidental preponderance of 'A' serial orders in Experiment IA. If the results of Experiment IA are excluded (Table IXA), then no difference whatever between serial orders exists.

C. *Memory Ability of Ss*

Zeigarnik and Marrow both found that the ratio of recalled interrupted to recalled completed tasks declined as the memory ability of the *Ss* increased. This is to be expected on a chance basis, since the more tasks recalled altogether the greater the chance of representation from both interrupted and completed tasks. If, therefore, an *S* recalls all 18 tasks, his ratio is bound to be 1.00; if he recalls 16, it can be no greater than 1.29 (9/7); if 14, no greater than 1.80 (9/5), etc.

We also found that the ratio of recalled interrupted to recalled completed tasks declined with increased memory ability of the *Ss* (Table X). We also see from this table that the distribution of total

TABLE X

Analysis of RI in Terms of Memory Ability of *Ss*

Experiment	No. Tasks Recalled	No. of *Ss*	Ratio RI/RT	Ave. RT
I	6–9	6	.66	10.70
	10–14	6	.63	
IA	6–9	7	.37	9.80
	10–14	5	.40	
II	6–8	6	.66	8.20
	9–12	5	.55	
III	6–10	11	.56	10.30
	11–14	12	.53	

number of tasks recalled (memory ability of *Ss*) is not unduly weighted either in the low recall or high recall direction under any experimental condition. This rules out the possibility that this factor is operating to produce our results.

Discussion

One final table (Table XI) throws light on the nature of the forces operating in the various experiments. In a sense, Experiments I, II and III can be considered as 'controls' for Experiment CW. An

TABLE XI

COMPARISON OF CW WITH ALL OTHER EXPERIMENTS

Experiment	N	Ave. RI	Ave. RC	Ave. RI / Ave. RC	% RI (Ave. RI/Ave. RT)	% RC (Ave. RC/Ave. RT)	Ave. RT
CW*	14	3.71	4.21	0.88	0.47	0.53	7.92
I	12	6.25	3.58	1.75	0.64	0.36	9.83
IA	12	3.75	6.00	0.63	0.38	0.62	9.75
II	11	4.91	3.27	1.50	0.60	0.40	8.18
III	23	5.60	4.70	1.20	0.54	0.46	10.30
I, II, III	46	5.60	4.07	1.38	0.58	0.42	9.67

* These results from Table I of preceding paper (5).

objective orientation existed in all these experiments, coöperation existed in one, and objective completion in another. Only in the CW Experiment were the tasks completed by the coöperating partner. If the results of the CW Experiment are placed side by side with the results of Experiments I, II and III combined, it is strikingly clear that the difference between CW and the others is in recall of *interrupted* tasks. The average number of completed tasks recalled is the *same*, the average number of interrupted tasks is greater in the combined 'controls' than in CW. Quite obviously, then, completion by the coöperating partner did effect a release of the task-completion tension-systems in the CW Experiment.

It is interesting, also, to see how clearly the recall ratio reflects the conditions prevailing in each of the several experiments. When the person works alone, and half the tasks he is doing are interrupted, then recall favors interrupted rather than completed tasks in the ratio of nearly 1.75 to 1 (Experiment I). When the person works in coöperation with someone else, and half the tasks are interrupted, interrupted tasks are again recalled better than completed, this time in the ratio of 1.50 to 1 (Experiment II). When the person works alone and half the tasks are interrupted and completed by someone else, interrupted tasks are recalled in the ratio of only 1.20 to 1—still favored slightly in recall, but not nearly to the extent found without external completion (Experiment III). When the person works in coöperation with someone else, and half the tasks are interrupted by the partner and completed by the partner, then interrupted (partner-completed) and completed (self-completed) tasks are equally recalled (Experiment CW).

All of these results depend upon the existence of task-orientation in work. That is, these results are obtained only if the *S*'s main concern seems to be to complete his work, rather than to enhance his ego. For, when the person is ego-oriented, and half the tasks (at which he is working alone) are interrupted, recall favors completed

rather than interrupted tasks (Experiment IA). Complicated questions of 'success' and 'failure' enter here, as soon as the task-goal becomes a means for the satisfaction of an ego-goal.

The implications of this experimental demonstration of the results of task and ego-orientation are quite far-reaching. Particularly do the results of the CW experiment and of Experiment III offer support to the thesis that man's motivation in work is often a direct function of the requirements of the task he has undertaken. Experiments CW and III make it apparent that, on certain occasions, man's selfish needs are so little a part of the motivational system which guides him that participation of his 'self' in a task is not even necessary for the achievement of his goal. The goal is reached when the task is done; the agency of doing need not be the self.

In this light, a revised concept of the 'ego' might be introduced. For purposes of clarity in the development of these experiments, we have limited the term 'ego' to the meaning, 'selfish' or 'egotistical.' Actually, we have been describing, by this restriction of meaning, the type of situation in which man's personal needs are so narrow that they encompass only what relates to himself. This is the type of behavior which prevailed in Experiment IA. In contrast, we have also described experimentally, some situations in which man's personal needs are broad enough to encompass the needs of others, needs arising from the world around him. One can think of rigid and narrow 'ego-boundaries' in contrast to flexible and broad 'ego-boundaries.' In the latter case, the 'ego-boundaries' may include the needs of other selves or 'egos,' of groups, of ideals.

Experimental study of the consequences of narrowing of ego-boundaries, that is, of ego-orientation in work, is also needed. The senior author has undertaken some experiments in this field which will shortly be reported.

Summary and Conclusions

1. The greater recall of interrupted tasks found by Zeigarnik and others depends upon the existence of task-orientation. When the person is ego-oriented, then recall favors the completed (ego-enhancing, 'success') tasks (Experiments I and IA).

2. Interrupted tasks are better recalled in coöperative as well as in isolated work (Experiment II).

3. Completion of a task by another person (objective completion) may be as satisfactory as personal completion, even when the person is working alone. This depends upon task-orientation and upon the nature of the task (Experiment III).

(Manuscript received January 17, 1944)

References

1. Abel, T. M. Neuro-circulatory reaction and the recall of unfinished and completed tasks. *J. Psychol.*, 1938, **6**, 377–383.
2. Adler, D., & Kounin, J. Some factors operating at the moment of resumption of interrupted tasks. *J. Psychol.*, 1939, **1**, 255–267.
3. Freud, S. Repression. *In: Collected Papers*, 1915, IV, 85–97.
4. Lewin, K., Lippitt, R., & Escalona, S. Studies in topological and vector psychology. I. *Univ. Iowa Stud. Child Welf.*, 16, No. 3.
5. Lewis, H. B. An experimental study of the role of the ego in work. I. The role of the ego in cooperative work. *J. exp. Psychol.*, 1944, **34**, 113–126.
6. Marrow, A. J. Goal tensions and recall. I and II. *J. gen. Psychol.*, 1938, **19**, 3–35; 37–64.
7. Rosenzweig, S. Preference in the repetition of successful and unsuccessful activities as a function of age and personality. *J. gen. Psychol.*, 1933, **42**, 423–441.
8. Rosenzweig, S., & Mason, G. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, **24**, 245–265.
9. Rosenzweig, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reaction to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.
10. Zeigarnik, B. On the recall of completed and incompleted tasks. *Psychol. Forsch.*, 1927, **9**, 1–85.

THE EFFECT OF DIFFERENT AMOUNTS OF REINFORCEMENT UPON THE ACQUISITION AND EXTINCTION OF A SIMPLE RUNNING RESPONSE

BY FREDERICK A. MOTE

University of Connecticut

Introduction

Graham and Gagné (3) have reported the results of a series of experiments concerned with the acquisition, extinction and spontaneous recovery of a simple running response in the white rat. The experimental situation required the animal to leave a starting box, traverse a short runway, and enter a food box where he received reinforcement. The measure of behavior employed was the 'latent period' of the animal's response, i.e., the time required for him to cross a mark on the runway a short distance in front of the starting box after the presentation of the conditioned stimulus (the opening of the starting-box door). Graham and Gagné present data for 15 acquisition trials and five extinction trials.

The experiment to be reported here provides information upon the course of acquisition and extinction for different amounts of reinforcement. Five groups of rats were given respectively 3, 6, 12, 18, and 24 reinforced acquisition trials followed by a series of extinction trials. The results indicate that at the final level of acquisition statistically significant differences in latent period values are found between the 3-trial group and the 12-, 18-, and 24-trial groups but that none of the other differences between groups is significant. In the extinction data the most significant feature to appear is the variation in the slopes of the extinction curves; in general, these curves tend to be steeper the greater the number of acquisition trials.

Apparatus and Procedure

The apparatus has been fully described elsewhere (1, p. 54). Except that provision was made to prevent the animals' climbing from the runway onto the starting and food boxes, it was similar to the apparatus used by Graham and Gagné.

The experimental animals were males about two months old, divided into five groups of 11 animals each. They are designated as Groups I, II, III, IV, and V, and were given, respectively, 3, 6, 12, 18, and 24 reinforced acquision trials. In all cases the acquisition trials were followed immediately by a series of non-reinforced extinction trials.

Each animal was given four days' preliminary training prior to the experimental session. The first day the 24-hour hungry rat was allowed to run about on a practice runway for 10 min. He was then transferred to one of the experimental boxes which was placed for five min. in the

XUM

starting-box position in the apparatus. The box and rat were then moved to the food-box position, the animal was allowed to eat for 20 min., and then was finally returned to the living cage. Next day the same procedure was repeated, employing the box not used on the previous day. On the last two days, while the rat was in the food-box position but before the 20 min. feeding, he was given three small pellets of food, one at a time, in order to insure his picking up and eating the reward pellets during the experimental session. At all times the food was Purina rabbit chow. During the preliminary training there was no opportunity for the animal to become habituated to the experimental runway, since the doors were kept closed.

The following experimental routine was carried out on the fifth day. After 10 min. on the practice runway the rat was transferred to one of the boxes, and box and rat were placed in the starting position. The other box with a small pellet of food (about 0.04 gm.) was placed in the food-box position and the door opened. After three min. the starting-box door was opened and a stop watch was started. When the rat left the box and crossed a mark four in. in front of the starting box door, the watch was stopped. This measure constituted the 'latent period.' The time was noted and the watch set back to zero. (If the animal turned back after crossing the mark, the watch was immediately re-started and the total time until the final crossing of the mark was recorded.) As soon as the animal entered the food box, the door was closed and the watch was started again. The starting-box door was closed, and at the end of the 30 sec. allowed for consuming the reward pellet the food box with the rat in it was moved back to the starting position and the other box (with a reward pellet) was placed in the food-box position. Thus, the boxes were used interchangeably so that the starting box of one trial became the food box of the next. Seventy-five sec. after the rat had entered the food box the starting-box door was opened and the next trial was begun. The time relationships involved in a single trial were: a fixed inter-trial interval of 75 sec., composed of 30 sec. in the food box and 30 to 35 sec. in the starting-box position (45 sec. minus the time required to move box and rat from one position to the other); and a latent period interval which varied according to the animals' response. Seventy-five sec. after the last acquisition trial the first of the series of extinction trials was begun.

Exactly the same procedure was followed during the extinction series, but since no food was given the rat entered an empty food box. The criterion of extinction was arbitrarily set as failure to leave the starting box within three minutes after the door was opened. If the animal left the box but turned back after crossing the four-in. mark the watch was started again and a total of three min. was taken as the indication of extinction. Preliminary work revealed that the number of trials required to reach this extinction criterion varied greatly from animal to animal; some required as few as three trials, others more than 100. Therefore a limit was set for the number of trials: if the rat failed to reach the criterion in 20 trials, his extinction series was discontinued.

Results

Acquisition.—The average data for all groups are presented in Table I. In Fig. 1 the course of acquisition is plotted as a function of the number of trials. The logarithm of the latent period (in sec.) at each trial has been used as the unit of measurement in order to conform to the practice which has been employed in studies involving this experimental situation. (For an explanation of the reason for this procedure see **3**, p. 258.)

Figure 1 shows that the average log latent period values are approximately the same for the first trial, ranging from 1.92 (83.2 sec.) for Group I (3 trials) to 2.11 (129 sec.) for Group II (6 trials). Following this first trial the latency values drop sharply to around 1.25 log units (17.8 sec.) for trial 2. These values for the different groups on the first two trials are so nearly the same that it is difficult to identify the individual group curves. From trial 2 onward the decline in latent period becomes progressively less and less with each

TABLE I

AVERAGE LOG LATENT PERIOD VALUES AND A.D. LOG LATENT PERIOD VALUES FOR ACQUISITION TRIALS AND THE FIRST 10 EXTINCTION TRIALS

Acquisition Trial	Group I		Group II		Group III		Group IV		Group V	
	Log L.P.	A.D. Log L.P.	Log L.P.	A.D. Log L.P.	Log L.P.	A.D. Log L.P.	Log L.P.	A.D. Log L.P.	Log L.P.	A.D. Log L.P.
1	1.92	.32	2.11	.24	1.96	.25	2.00	.37	1.97	.27
2	1.23	.22	1.28	.18	1.23	.17	1.18	.22	1.33	.22
3	1.04	.27	.92	.20	1.13	.25	1.03	.35	1.04	.23
4			.82	.24	1.12	.20	1.15	.36	1.02	.34
5			.79	.26	1.18	.31	.99	.46	.90	.37
6			.94	.27	.97	.22	.96	.34	1.95	.25
7					.95	.37	.96	.35	.86	.27
8					.99	.44	.78	.20	.98	.24
9					1.10	.44	.79	.35	.97	.35
10					.92	.33	.79	.23	.84	.21
11					.96	.38	.76	.36	.74	.24
12					.89	.25	.87	.29	.74	.27
13							.78	.35	.68	.16
14							.79	.44	.73	.19
15							.79	.33	.66	.36
16							.83	.32	.59	.20
17							.68	.44	.90	.33
18							.74	.37	.82	.21
19									.89	.25
20									.81	.27
21									.82	.28
22									.74	.22
23									.67	.24
24									.74	.29
Extinction Trial										
1	1.03	.33	.89	.37	.89	.32	.72	.28	.75	.39
2	1.12	.20	1.07	.33	.98	.33	.78	.35	.77	.38
3	1.22	.43	1.09	.42	1.01	.32	.94	.41	.95	.34
4	1.18	.43	1.04	.37	.96	.39	1.22	.36	1.14	.53
5	1.31	.44	1.28	.26	1.15	.48	1.32	.38	1.28	.32
6	1.26	.39	1.20	.42	1.38	.54	1.40	.39	1.17	.32
7	1.45	.30	1.16	.36	1.40	.36	1.76	.29	1.34	.40
8	1.40	.40	1.28	.37	1.45	.44	1.46	.32	1.12	.23
9	1.33	.28	1.26	.27	1.28	.45	1.34	.40	1.58	.26
10	1.38	.42	1.48	.32	1.41	.43	1.21	.30	1.62	.21

succeeding trial, and is marked by fluctuations from trial to trial. At about trial 15 the curves appear to be approaching a relatively steady value of approximately 0.75 log unit (5.6 sec.).

The curve which most clearly shows the levelling off at about 0.75 log unit is the curve for Group IV (18 trials). Group III (12

trials) has not reached this point by the end of its acquisition series, whereas Group V (24 trials) shows, between trials 16 and 17, a marked rise which is sustained for several trials. If the shift from a value of 0.59 to 0.90 log unit for this group were followed by an immediate return to about the former value it might be dismissed as a chance fluctuation involving only a single trial. However, the curve only gradually descends from the 0.90 figure at trial 17, and it is not until trial 22 that a value of 0.74 log unit is reached, a level which this group had attained on trial 11. Inspection of the data for

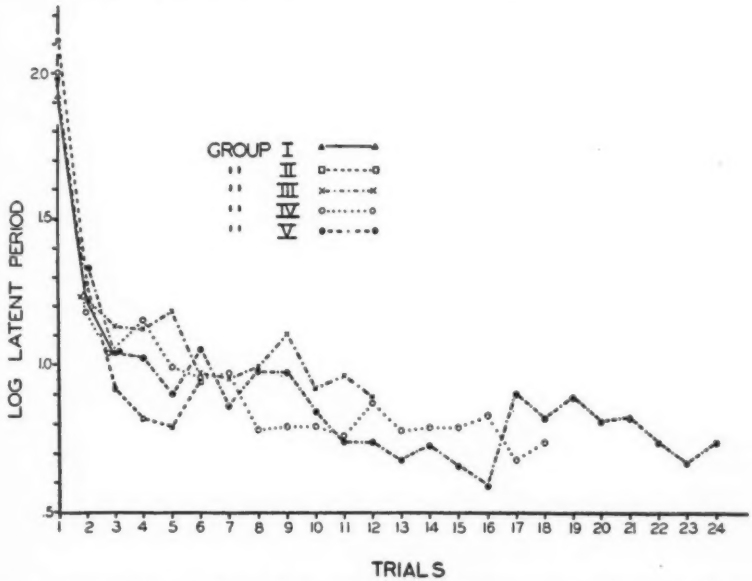


FIG. 1. Average latency values during acquisition for all five groups. Log latent period values obtained from 11 animals contribute to each point.

the individual animals comprising this 24-trial group reveals that four of the 11 are responsible for the greater part of the increased average values from trial 17 onward. It is not likely that satiation is responsible for this rise, as it has been demonstrated that this factor does not affect the behavior in this situation until about five gm. of food have been consumed; by the 17th trial these animals had received only a fraction of this amount. The factor of fatigue is also unlikely, as preliminary work on this problem and the results of a previous study (5) show that rats in this experimental situation frequently traverse the runway 40–50 times with roughly constant latent period and no apparent signs of fatigue.

In this experiment it was possible to combine the data so as to furnish information upon the course of acquisition for more animals than the 11 comprising each group. Data on 55 rats are available for 3 acquisition trials, 44 for 6 trials, 33 for 12 trials, and 22 for 18 trials. The mean log latent period values for each group are only slightly changed when these additional cases are added. For example, in Group I the mean values are 1.92, 1.23, and 1.04 log units for trials 1, 2, and 3 respectively. Adding the other 44 cases changes the values to 1.99, 1.22, and 1.03. For the remaining groups changes of similarly slight magnitude are obtained; only eight latent period values out of a possible 61 are changed more than 0.10 log unit, the greatest shift found being 0.21 log unit. The changes in the average deviations of the log latent period values are correspondingly slight.

The data obtained when more cases are added indicate that, insofar as the mean values from trial to trial during acquisition are concerned, confidence can be placed in the results obtained from the 11 animals comprising each group. An exception of this statement may be necessary when the last few trials of the Group V (24 trials) animals are considered. In a previous paragraph reference was made to the change in the course of acquisition for this group. Although there are no additional data for the complete course of acquisition for this group (i.e., all 24 trials) there are additional data from the 11 animals of Group IV covering the first 18 trials. When only the 11 rats of Group V are considered, the values are 0.59, 0.90, and 0.82 for trials 16, 17, and 18 respectively. After the Group IV data are included so as to furnish information from 22 cases in all, the values become 0.71, 0.79, and 0.78. Thus, the variability in Group V (chiefly brought about by four aberrant cases) has resulted in too low a value for trial 16 and values too high at trial 17 and 18; doubling the number of cases tends to smooth the curve.

In another paper (**5**) it has been shown that a value of about 1.61 log units may be used as the basal value from which to compute changes in latent period in this and similar problems employing the same response and experimental procedure. This figure was obtained as the average log latent period over 10 trials for a group of animals subjected to the same preliminary training and experimental routine as that of the present study except that reinforcement was never given. Thus, 1.61 log units is the average value which may be taken to represent the rat's latent period when only the 'exploratory' drive is involved in the situation. Some conception of the magnitude of the change in latent period occurring during the course of acquisition in the present experiment can be obtained by comparing this figure with the mean log latent period values at the final level of acquisition for each group. (The first trial of the extinction series

is taken as the final level of acquisition since this trial, insofar as the latent period is concerned, has not had the opportunity to be modified by the lack of reinforcement but shows only the effect of previous reinforcements.) The comparison is made between the latent period values for each group at this first non-reinforced trial and 1.61 log units. For Group I the magnitude of the change is of the order of 3.8 to 1 (1.61 log units = 40.7 sec., 1.03 log units = 10.7 sec.); for Group II, 4 to 1; Group III, 5.8 to 1; Group IV, 6.5 to 1; and Group V, 7 to 1.

Statistical calculations of the data were made in order to determine the significance of the difference between mean log latent period values after the different amounts of reinforcement. The values derived from the combined data were used. The analysis shows that at the first trial in the extinction series the difference between Groups I and II (3 versus 6 trials) is at the 10–20 percent level of significance. The differences between Group I and the remaining three groups (12, 18, and 24 trials) are all significant at the level of two percent or less. None of the differences between the other groups is significant; for example, between Groups II and III (6 versus 12 trials) the difference is at the 10 percent level and between Group II and Groups IV and V (6 versus 18 and 24 trials respectively) at the 20 percent level. Thus 6 reinforcements do not significantly affect the mean log latent period as compared with 3 reinforcements, nor do 12 as compared with 6, but 12 reinforcements do significantly reduce the latent period as compared with 3. Beyond 6 reinforcements there is no statistically significant change detectable.

Extinction. The extinction data for the five groups are plotted in Fig. 2. Only the first 10 trial values are shown in both Table I and Fig. 2, as the numbers of animals continuing to respond beyond 10 trials rapidly diminished; in no case did all 11 animals of a group complete 10 extinction trials.

The general trend in the direction of a rapid increase in latent period is easily seen in the curves representing the course of extinction. Although the curves show marked irregularities, all groups are responding with a considerable longer latent period by the 10th trial than at the beginning of the extinction series. In Finger's data (1) the group given the greater number of reinforced acquisition trials showed more rapid approach toward the criterion of extinction; the slope of their curve was steeper. In general, the same trend is apparent in Fig. 2. If the log latent period value of the first trial of the extinction series is subtracted from the averages of the 9th and 10th trials in the series the following values are obtained: Group I = 0.32 log unit (average of trials 9 and 10 = 1.35 log units = 22.4 sec.; trial 1 = 1.03 log units = 10.7 sec.); Group II = 0.46; Group III

= 0.46; Group IV = 0.56; and Group V = 0.85. These figures indicate that the greater the number of reinforced acquisition trials, the steeper becomes the slope of the extinction curve; i.e., the less the resistance to extinction if the measure employed is the rate of approach toward the criterion of extinction.

Fig. 2 shows that on the fifth trial in the extinction series the different groups all have about the same mean log latent period value. Beyond this point the number of animals in each group is smaller (as few as seven by the 10th trial for Group V) and the use of an average becomes questionable. The statistical calculations are

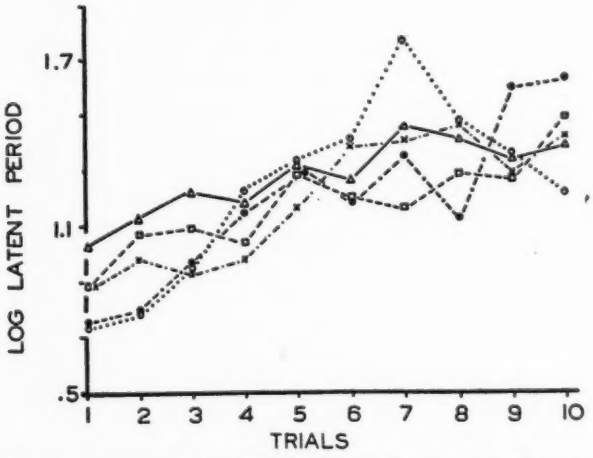


FIG. 2. Average latency values during extinction for all five groups. Data for 11 animals contribute to trials 1–4 for all groups; after the fourth trial *N* is variable. The group designations for the curves are the same as in Fig. 1.

therefore based upon the latent period values at the fifth trial. The significance of the difference between trial 1 and trial 5 in the extinction series for Group I is at the 20–30 percent level; that for Group II at the 2–5 percent level; that for Group III at the 20–20 percent level; and that for Groups IV and V at about the one percent level. Thus, except for Group III, as the number of acquisition trials increases, the significance of the difference between the first trial in the extinction series and some selected point later in extinction tends to increase.

DISCUSSION

The outstanding feature to appear in the acquisition data is the levelling off in the latent period values which occurs from about the 12th trial onward. This is also found in several other reports which

have dealt with the problem of acquisition in this experimental situation (1, 3). In a more recent study (6) the same result was obtained when the animals were given 32 reinforced acquisition trials. In the section on experimental results it was noted that no statistically significant differences were obtained in the present experiment when comparisons were made between the final levels of acquisition for the groups receiving 6, 12, 18, and 24 trials. The difference between Group I (3 trials) and Group 11 (6 trials) was not significant but the differences between the 3-trial group and the 12-, 18-, and 24-trial groups were all significant. Such comparisons indicate that at some amount between 6 and 12 reinforcements a level of response is reached beyond which the addition of more reinforcement is ineffective in bringing about a further decrease in latent period. In other words, insofar as this acquisition measure is concerned the excitatory effect of, e.g., 9 reinforcements is just as effective as 12, 18, or 24.

This conclusion does not mean that no effect of additional reinforcement exists, but merely that this statistical analysis of the latent period measure fails to reveal it at the rather exacting statistical significance level of one percent or less. There is a trend in the direction of greater effectiveness of reinforcement as the number of acquisition trials is increased. It may be that some measure other than the latent period would reveal more definitely the influence of the additional reinforcements.

The 75-sec. inter-trial interval used in this experiment is a rate which falls into the category of 'massed' practice according to the results of an experiment upon the effect of varying the inter-trial interval (2). An hypothesis (3, 4) couched in terms of the factors of excitation and inhibition as they affect behavior under the conditions of massed practice and which can be applied to explain the flattening of the acquisition curve is the following: At each reinforced trial an increment of inhibition is added as well as an excitatory increment. Between trials the more 'labile' inhibition dies away and the longer the interval the more nearly is it completely dissipated before the next trial. With short intervals (i.e., massed practice) the inhibitory component is only partially dissipated from one trial to the next. The net result is that inhibition accumulates at such a rate relative to the accumulation of excitation that it effectively balances the excitatory effect of the latter, and the rate of acquisition approaches a constant value. The failure to obtain a further decrease in latent period values which occurs in this experiment about the 12th trial is therefore to be expected according to such an hypothesis.

Another possible explanation for the levelling off in latent period values may be the following: At any trial after the first in the acquisition series the behavior of the animal at the time the starting-box

door is opened may be assumed to be subject to these response tendencies: (1) The tendency to traverse the runway toward the food box due both to previous reinforcement and the 'exploratory drive.' (2) The tendency to remain inside the starting box wherein reinforcement has just been obtained when it was in the food-box position. At about the 12th trial these opposed tendencies reach an equilibrium, so that the positive effect of the first toward evoking more and more prompt traversal of the runway at each succeeding trial is balanced by the second, which has the effect of keeping the rat in the starting box. The orderly decrease in the latent period is therefore no longer obtained and the net result is the flattening of the acquisition curve. The tentative nature of this explanation should be emphasized as, so far as is known, there are no data which allow a precise and quantified exposition of such an hypothesis.

The outstanding feature to appear in the extinction data is that the slope of the extinction curves becomes steeper as the number of trials in the acquisition series increases. Thus the 6-trial curve is steeper than that for 3 trials while the curves for 18 and 24 trials are steeper than the other two. The curve for the 12-trial group is anomalous in this respect, since it parallels the curve for the 3-trial group. The curves for 18 and 24 trials have the same slope.

On an *a priori* basis it would probably be predicted that the curves would have a variation in slope which in general would be opposite to that actually obtained. That is, the curve for 3 trials would show the steepest slope (indicating least resistance to extinction) while for the groups with more acquisition trials, and therefore more reinforcement, the curves would exhibit flatter slopes.

In an experiment similar to the present one wherein one group of rats was given 16 reinforced acquisition trials and another group 8, Finger (1) obtained a like result, in that the former group showed less extinction resistance than the latter. The same relationship is found in this study between the 6- and 3-trial groups and the 24- and 12-trial groups. It does not hold between the 12- and 6-trial groups. Of the several hypotheses which Finger considers with respect to this phenomenon, one is concerned with the exploratory drive. The argument involving exploratory drive postulates that the greater the number of trials during the acquisition series the more this drive is satisfied. Thus, when the experimentally acquired response is undergoing extinction, the animals with more trials in their acquisition series will show more rapid extinction because their behavior is primarily influenced by the reinforced tendency to traverse the runway and relatively less influenced by any remaining unextinguished exploratory drive. For animals with few trials in acquisition the behavior during extinction is a function of two response

tendencies: the excitatory effect established during reinforced training and the remaining exploratory drive tendency; therefore these animals might be expected to show less rapid extinction.

It is now known (5) that animals will respond for an average of 16 times in this experimental situation when no food reinforcement is present. There is thus no question but that the exploratory drive could still have been present in the 3-, 6-, and 12-trial groups at the end of their acquisition series. It has presumably become extinguished in the 18- and 24-trial groups. (This is true when extinction is carried to the criterion of three min. of non-response, but its presence has been demonstrated up to an average of 28 trials when a five-min. criterion was used.) From about the fourth trial onward in the non-reinforced situation, the exploratory drive remains approximately constant at the previously mentioned value of 1.61 log units. When these facts are applied to the data of the present experiment it means that insofar as we have any measure of the exploratory drive, it exists in all the groups up to about 16 trials and to the same magnitude in the 6 and 12 trial groups at the end of their acquisition series. Thus, the drive to explore the apparatus can be invoked to help account for the differences found in the extinction data between the 3-trial group and the other groups and between the 18- and 24-trial groups and the 12-trial group. It will not explain the differences between the groups having received 12, 18, and 24 trials and the 6-trial group.

Another aspect of the hypothesis employing the interplay of excitation and inhibition which was used to give a possible reason for the levelling of the acquisition curve may be used to explain the results found in the extinction data. Extinction, according to this view, consists in both the 'draining off' of excitation by repeated non-reinforced elicitation of the acquired response and the active suppression of excitation by an inhibitory effect. With respect to the extinction phenomena under discussion this explanation would apply as follows: At the final level of acquisition for the 12-, 18-, and 24-trial groups the amount of inhibition relative to the amount of excitation is such that the response time is approximately constant. This indicates that the inhibitory effect is the dominant factor in the situation. If we make the reasonable assumption that its rate of dissipation is proportional to the amount of inhibition present, the more rapid extinction of the groups with more acquisition trials is a necessary consequence.

Summary and Conclusions

1. Five groups of 11 rats each were given 3, 6, 12, 18, and 24 reinforced acquisition trials in an experimental situation requiring

them to leave a starting box, traverse a short runway, and enter a food box. The measure of behavior employed was the latent period of response, i.e., the time elapsing between the opening of the starting-box door and the crossing of a mark on the runway four in. in front of the door. A series of extinction trials was begun immediately after the last acquisition trial. Curves for the acquisition and extinction data of all groups are presented. In these figures the logarithm of the latent period (in sec.) is plotted as a function of the number of trials.

2. In the acquisition data statistically significant differences (at the level of two percent or better) were found between the group having 3 trials and those having 12, 18, and 24 trials. None of the other group differences is significant at this level, although there is a trend in the data showing that the greater the number acquisition trials the shorter the latent period at the final level of acquisition.

3. In the extinction data the most noticeable feature was that, in general, the slope of the extinction curve became steeper the greater the number of acquisition trials. Thus, it would appear that the greater the amount of reinforcement the less the resistance to extinction.

4. The results are discussed in the light of some possible explanations regarding behavior in this experimental situation.

(Manuscript received February 23, 1944)

Bibliography

1. Finger, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
2. Gagné, R. M. The effect of spacing of trials on the acquisition and extinction of a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 201–216.
3. Graham, C. H., & Gagné, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
4. Hovland, C. I. 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. nat. Acad. Sci.*, 1936, **22**, 430–433.
5. Mote, F. H., Jr., & Finger, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, 1942, **31**, 57–68.
6. Mote, F. A., Jr., & Finger, F. W. The retention of a simple running response after varying amounts of reinforcement. *J. exp. Psychol.*, 1943, **33**, 317–322.

GUTHRIE'S THEORY OF LEARNING: A SECOND EXPERIMENT

BY JOHN P. SEWARD, JANE B. DILL, AND MILDRED A. HOLLAND

Connecticut College

Problem

According to Guthrie, in a trial-and-error situation each response becomes associated with the stimuli active at the time. At the same time it dissociates the preceding response from those stimuli. But the successful response removes the disturbing stimuli from the situation and therefore cannot be dissociated from them by any subsequent reaction. The successful response is selected, not because it is followed by reward, but simply because it comes last (**6**).

Guthrie's theory is especially welcome to those who have difficulty in fitting the concepts of 'effect' and 'reinforcement' into their scheme of things. It offers a convenient way to avoid the clumsy notion of a connection being somehow strengthened by a subsequent event quite different in kind.[1] The senior author has already relied on it heavily in attempting to formulate a theory of maze discrimination (**18**). It therefore becomes a matter of some moment to provide Guthrie's finality theory with empirical support.

In a previous experiment to this end (**17**) the senior author attempted to compare the effect of food-reward on the strength of a response with that of finality alone. Two groups of hungry rats were placed in a bar-pressing set-up. Upon pressing the bar one group was rewarded by food, the other was merely removed from the box. Both groups showed some learning, the former more than the latter. But the appearance of anticipatory responses in both groups suggested two things. First, Guthrie's description of selective learning was probably oversimplified; and secondly, the experiment had failed to isolate the factor of finality.

The present study represents another attempt to provide a crucial test of the finality theory. In this case the approach was quite different. Instead of testing its applicability to trial-and-error learning, we took our departure from the proposition that when a response is associated with a stimulus, any previous response to that stimulus

[1] Previous reference to this aspect of Guthrie's doctrine as the finality theory *of reinforcement* (**17**) was, strictly speaking, inaccurate. Guthrie himself uses the term 'reinforcement' correctly in connection with the function of emotion (**8**). To explain selection in trial-and-error learning he considers the concept unnecessary.

is dissociated from it. The deduction follows that if a stimulus evokes first one and then another response, other things being equal, it will thereafter tend to arouse the second response rather than the first.

Evidence bearing directly on this deduction is surprisingly scarce. Experiments on negative transfer and retroactive inhibition frequently involve learning two different responses to the same stimulus (**14**), but they seldom if ever provide a measure of the relative strengths of the two responses after equal practice. Studies of habit reversal (**1, 3, 5, 9, 13, 21**), which constitute a special case of the negative transfer experiment, are on the whole open to the same objection and prove to be hopelessly equivocal with respect to our problem. More pertinent is the experiment of Kellogg and Wolf (**12**), who, working with dogs, made a buzzer the conditioned stimulus to retraction first of the right forepaw, then of the left. In spite of much longer training on the first CR, retention tests after one to six months consistently evoked the second. The authors concluded that "the movements most recently conditioned will be the movements reproduced after an interval," a conclusion clearly in line with the theorem here under consideration. As the authors suggested, however, successive conditioning usually involves extinction of the first CR along with setting up the second. Indeed, when it does *not*, as in 'ambiguous conditioning,' *both* CRs are typically retained. Until extinction itself has been proved to consist of the displacement of one response by another (**10**), we are guilty of begging the question if we invoke it as evidence.

The one experiment most closely resembling ours is Calkins's 1896 study of the conditions of association (**4**). Her purpose was to measure the relative effectiveness of frequency, recency, primacy, and vividness. She presented a series of color-number pairs within which one color was paired with two numbers, one 'normal,' the other emphasized by one of the above-mentioned factors. In a retention test immediately following, she presented the colors in a changed order and determined the percentage of times each of the two alternative responses was recalled. Of greatest interest to us were her data on primacy and recency. With visual presentation, for example, when one of the two numbers came first in the series it was recalled in 36.5 percent of cases, as compared with 29.5 percent for the 'normal' response. When one of the two numbers came last the corresponding figures were 53.7 percent and 25.7 percent. From these and other data Calkins concluded that recency was much more effective than primacy under these conditions.

At first glance Calkins's findings appear to bear out Guthrie's theorem. Closer inspection shows that they do so only indirectly

if at all. For Calkins was concerned with firstness and lastness—as compared with middleness—in a series, while Guthrie's theorem concerns the firstness and lastness of the two responses solely with respect to each other. To measure the influence of this factor, that of the former—intraserial position—must be held constant.

The present experiment was therefore devised to measure the relative strengths of the first and second responses to a stimulus.

Method

The process of measurement involved essentially three steps: (1) associating S with R_1, (2) associating S with R_2, (3) presenting S and measuring the relative tendencies for R_1 and R_2 to occur. Our problem also presented four conditions which had to be met:

1. The two responses must be equally rewarded.
2. The two responses must be equally strong before the experiment.
3. The two responses must be equally easy to learn and recall.
4. Some learning must take place, but not enough so that both responses can be recalled as correct and a deliberate choice made between them.

To meet these requirements a memory experiment was devised in the form of a serial multiple choice test. Within the series one stimulus was presented twice, each time paired with a different response. In the recall test this stimulus was given just once to determine whether the first or second response would be made.

S was seated at a table before a row of 12 push-buttons. In an upright panel facing him was a rectangular window 4 in. × 2½ in. and below it a small round opening. *S*'s instructions were essentially as follows:

"You will be shown a series of colored cards in the upper window. For each color you are to press buttons until you find the one which turns on the light in the round opening. The buttons do not have to be pressed in any particular order. This is a memory experiment."

Ten colored papers of the Milton Bradley series were exposed once each, long enough for *S* to find the correct button. Before each exposure *E* put the button in circuit by a selector switch. Without informing *S*, one color was presented twice, once in the fifth position and again in the eighth, making 11 exposures altogether. After an interval of approximately five min. *S* was instructed as follows:

"Now you will be shown the same colors in a different order and will be given six chances to find the correct button for each one."

E then presented the 10 colors and recorded each button pressed by *S*. During the test the two buttons which had been paired with the repeated color were connected so that either one would operate the light.

The colors in order (Milton Bradley designations), with the corresponding buttons numbered from *S*'s left to right, were as follows:

Presentation Color	Button	Test Color
1. GY	10	1. B (t2)
2. RV (s1)*	12	2. RV (s1)
3. RO	3	3. OR
4. B (t2)*	11	4. GY
5. BG	5 or 8	5. B (s1)
6. B (s1)	4	6. V (t2)
7. OR	2	7. G
8. BG	8 or 5	8. RO
9. G	1	9. O (t2)
10. O (t2)	6	10. BG
11. V (t2)	7	

* s—shade; t—tint.

How did this procedure control the factors mentioned in the four conditions outlined above? The first condition, that the two responses be equally rewarded, was met at least partially by having both responses turn on the light. How adequately rewards as a whole were equated, however, will be taken up again below. (See Results.)

The second factor, the relative initial strengths of the two responses, was controlled by dividing the *Ss* into two groups. In one group the first correct response to BG was button 5 and the second was button 8; in the other group the order of the two buttons was reversed. These buttons were chosen as symmetrically placed in the row of 12 and presumably not too easy to identify. The effect of any preference for one—say No. 8, on the right—was nullified when the groups were combined.

The third factor, the relative ease of learning and recalling the first and second responses, was a function of intraserial position. The fifth and eighth positions in order of presentation were chosen as presumably of about equal difficulty (14). To confirm this presumption, however, a control group was run through the same procedure except that instead of repeating the fifth color in position 8 a new color was introduced at that point. Further, the control group was divided in order to counteract inequalities either in stimuli or responses. In one half, the fifth and eighth S-R pairs of the presentation series were BG-5 and Bk-8; in the other half, they were Bk-8 and BG-5, respectively. The fifth and eighth colors appeared in the eighth and eleventh positions, respectively, on the test. The control group thus provided a measure of the relative ease of learning at the two critical positions in the series when the stimuli were not identical. In a further attempt to minimize inequality of recall due to serial position, a five-min. interval was introduced before the test and the critical stimulus was placed last in the test series. These measures were aimed to provide enough time for unequal perseverations to subside.

The fourth condition, that the two responses should be learned but not clearly discriminated, was responsible for the form of the experiment. All the other colors and buttons merely served to increase the difficulty of learning and to disguise the fact that two responses had actually been associated with the same stimulus. To discover how successfully the second objective had been achieved, each *S* was questioned at the close of the experiment. He was first asked:

1. "In the first series, did you notice if any color was repeated? If so, which?"

If the correct color was identified, he was then asked:

2. "When it appeared the second time, did you recall the button that had been correct before?"

3. "When it appeared on the test did you recall both responses? If so, on what basis did you make your choice?"

The *Ss* were 170 women college students, 55 in each of the two experimental groups and 60 in the control group.

Results

The crucial question is: When the repeated color appeared on the test, how many *Ss* responded by pressing the button which had been correct the first time and how many by the second? The results on this point are summarized below:

Number of *Ss* making first response:	58
Number of *Ss* making second response:	42
Number of *Ss* making neither response:	10

These results fail to confirm our expectation, derived from the finality theory, that the last response to a stimulus should be preferred. If anything, there appears to be a tendency in the opposite direction. This tendency is not statistically reliable. Disregarding the 10 *Ss* who failed to make either response, we find that our obtained percentages deviate from chance by a ratio of 1.6, which is

significant only at the 10 percent level of confidence. But even a chance distribution is negative to our hypothesis.

Before rejecting the hypothesis let us examine other possible interpretations. First, if the fifth response of the series were easier to learn than the eighth, this fact might mask the finality effect. It was to check such a possibility that we introduced the control group, with the following results:

Percent of control *Ss* making correct response to 5th stimulus: 78
Percent of control *Ss* making correct response to 8th stimulus: 80

Evidently serial position cannot be held accountable for our results.

As a second possibility, chance results might simply indicate the absence of learning. If, in our zeal to complicate the situation, we had prevented our *Ss* from associating *either* response with the stimulus to a measurable degree, the experiment would have failed as a test of *any* theory. We are therefore compelled to furnish proof that learning actually took place at the critical points in the series. It will be recalled that *Ss* were allowed to try six buttons in response to each color on the test. Fortunately a record was kept of all test responses by each *S*. It was therefore possible to find the number of *Ss* taking one, two, three, or more trials to find the correct button and to compare the distribution with that to be expected by chance. The results for the critical stimulus are shown in Table I. The ob-

TABLE I

FREQUENCY DISTRIBUTION OF *Ss* ACCORDING TO NUMBER OF TRIALS TAKEN ON TEST TO PRESS BUTTON 5 OR 8, AS OBTAINED AND AS EXPECTED BY CHANCE

	Number of Trials							
	1	2	3	4	5	6	>6	Total
Obtained.....	24	25	22	21	4	4	10	110
Expected.....	18.3	16.7	15.0	13.3	11.7	10.0	25.0	110

$\chi^2 = 31.3$.

tained distribution is decidedly skewed toward the smaller numbers of trials, showing a performance superior to chance expectation. The χ^2 of 31.3 shows this difference to be completely dependable.

It may still be objected that better-than-chance performance does not necessarily indicate learning. For the repeated stimulus, it will be recalled, was reserved till the end of the test series. By that time a number of buttons would have been 'eliminated,' thus narrowing the range of choice to something considerably less than 12. Such a factor may indeed have operated. If it were the only one it should appear in the form of a gradual improvement in performance from

beginning to end of the test series, starting approximately at a chance level. Such a gradient should be reflected in the percentage of *S*s responding correctly to successive stimuli. These are presented in Table II. Inspection of the table shows that except for the extreme

TABLE II

PERCENTAGE OF *S*s RESPONDING CORRECTLY TO SUCCESSIVE STIMULI ON TEST

Group	Order of Stimuli										
	1	2	3	4	5	6	7	8	9	10	11
Experimental	58	86	78	82	66	82	79	69	94		
Control.....	65	78	77	78	73	77	68	78	62	82	80

positions no gradient appears. Moreover, with few exceptions, all the percentages are significantly above the 50 percent to be expected by chance.[2] It is unlikely, then, that the elimination process played a preponderant role in the *S*s' performance.

The evidence suggests strongly that some learning took place. But there is still a third possibility; namely, that *too much* learning might also defeat the purpose of the experiment. If some *S*s were aware of the repeated color and recalled both responses, they might be able to offset the associative factors by deliberate choice. To investigate this possibility we analyzed the replies to the questions asked at the close of each session. Of 108 *S*s (two were not questioned), 40 thought that BG had been repeated in the presentation; the other 68 did not realize it. Of the 40 who recognized BG the second time, 24 failed to notice that the correct buttons were different; 11 were aware of the difference but unable to recall both responses; only three made a deliberate choice. Of these, two chose the first response and one chose the second.[3] Obviously the results cannot be attributed to over-learning.

A fourth possibility was revealed when we raised again the problem of equality of reward and realized that the *consequences* of the two competing associations were not completely controlled. As already pointed out, both responses to BG were rewarded once by the appearance of the light. But whether or not one was made *without* being rewarded, when the other was correct, was left uncontrolled and unrecorded. Theoretically such an occurrence may have weakened the tendency to make that response to that stimulus just as being rewarded may have strengthened it. Whether it occurred more

[2] For the experimental group 60 percent represents a deviation significant at the 1 percent level; for the control group, 66 percent.

[3] The two remaining *S*s were unclassified.

frequently in connection with the first or the second response is unknown. Two opposite assumptions are conceivable. (1) On the one hand, it seems reasonable to suppose that the first response would be more often unrewarded than the second. For after the fifth button, say, had been established as the correct response to BG, it should be more likely to be pressed when BG appeared again. This supposition applies particularly to the *S*s who recognized the color. (2) On the other hand, if BG was not recognized, *S*'s mental set to *avoid* buttons already 'used up' would tend to keep him from making the same response again. In such cases the first response would be *less* often unrewarded than the second.

A rough check on these two assumptions is available in our data. As to the first: of the 40 *S*s who recognized BG as repeated, two failed to make either correct response; the other 38 were equally divided between the first and the second. In these cases inequality of reward should have favored the second response. As to the second assumption: of the 62 *S*s who did *not* recognize BG and who responded correctly on the test, 63 percent made the first response and 37 percent made the second. These percentages deviate from chance by a ratio of 2.03, which is significant at the 5 percent level of confidence. To this extent the assumption is upheld. Such arguments, of course, cannot take the place of data on the actual numbers of unrewarded responses in these groups.

Instead of accepting the last-named factor as responsible for our results, we prefer to suggest still a fifth interpretation. Our *S*s, it will be recalled, had to pick out the correct one (or two) from 12 buttons in response to each of 10 colors. This task involved the discrimination of both colors and buttons from one another. In one presentation the *S*s, for the most part, had reached a stage at which neither stimulus- nor response-discrimination was complete. On the stimulus side, some *S*s reported that other colors than BG seemed to have been repeated. On the response side, Table I shows that discrimination was by no means perfect. It was noticed that the two ends of the row of buttons were occasionally confused. *S*, for example, would try first the second from the left and then the second from the right. As noted above, most of the *S*s did not discriminate between the two critical responses. It seems likely, then, that buttons 5 and 8, chosen for their difficulty and equality, were so well chosen as to be practically equivalent. The principle of response generalization seems applicable here, according to which a stimulus becomes connected not only with the particular reaction made but with other related reactions (14, p. 421). In the present case BG, on its first appearance, would become associated with the correct button, say 5, and to some extent with other buttons, among them 8. On

its second appearance its association with 8 would be strengthened and at the same time the connection with 5 would be confirmed. Such a process of mutual reinforcement would tend to equalize the two connections, thus explaining why neither first nor second response had a clear advantage over the other. Its relation to Guthrie's theory will be indicated below.

Discussion

When a theorem is upheld by experiment, that theorem and all those propositions from which it was (or could have been) derived receive some measure of empirical support. At the same time the probability of all predictions based on the theorem is increased. When a theorem is experimentally refuted, however, the implications are not so far-reaching. The fault may lie in any of the postulates and theorems giving rise to the one in question but is not equally shared by all. Similarly, not all deductions from the theorem are weakened, but only those which cannot be otherwise derived. When an experiment yields negative results it therefore becomes especially important to determine the exact point of weakness in the theoretical foundations and the degree to which the superstructure is imperiled.

In the present case, for example, we have failed to confirm a prediction based on Guthrie's principles of learning. If we grant that the experimental conditions constituted a fair test of the prediction, is our confidence in his whole approach necessarily shaken? If not, where does the weakness lie? And to what extent do further derivations, such as his interpretation of selective learning, suffer?

In answering such questions it is a distinct advantage if the theory in question is presented as a logical system. Guthrie has recently set forth some of the basic assumptions underlying his position and drawn from them a number of deductions (**6**). His argument, however, emphasized the *fixation* of movements and included the *selection* of the successful response only in a somewhat incidental manner.

> "This prediction of a repetition in movement is complicated by another consideration. If . . . the cat establishes a second association with some repeated situation, and is led (through fatigue or refractoriness or pain or other cause) to respond differently, this will become its new response in that situation and there is evident a short-circuiting in its behavior" (**6**, p. 143).

This difference in emphasis was no doubt unintentional, since elsewhere in the same article Guthrie was most explicit in dealing with the function of the successful movement. We feel justified, therefore, in the interest of simplicity, in attempting a rudimentary recasting of that portion of his theory which is most relevant to our problem:

Definition 1: A stimulus which tends to arouse a response is said to *condition* that response.

Definition 2: A stimulus which tends to prevent a response is said to *inhibit* that response.

Definition 3: A persistent stimulus situation which keeps an organism disturbed and excited is known as a *problem.*

> "Problems are persistent stimulus situations of such a nature that they keep the animal or person disturbed and excited . . ." (7, p. 96).

Definition 4: A response which removes a problem is known as a *successful response.*

> ". . . A successful act may be defined as an act that terminates the persisting stimulation responsible for activity and excitement" (6, p. 140).

Postulate 1: "Stimuli which are acting at the time of a response become conditioners of that response" (8, p. 66).

Postulate 2: In becoming conditioners of a second response stimuli become inhibitors of the first.

> "A stimulus which has in the past contributed to one set of movements may cease to do this and may act as an inhibitor of such movements when it has been associated with another response" (8, p. 69).

Theorem 1: A stimulus which is acting while two or more responses are made will, on its recurrence, condition only the final response.

> "It can be reasonably argued that of two responses associated with a given substitute stimulus the second is always the one conditioned by that stimulus" (8, p. 97).
>
> "(In) simple conditioning . . . the last conditioning probably always prevails" (8, p. 102).

The theorem can be seen to follow directly from Definitions 1 and 2 and Postulates 1 and 2.

Corollary 1: A recurring problem will condition the successful response.

> "The next time that the disturbers are present they will tend to call out, by virtue of their last association, the act that removed them" (7, p. 98).

The corollary follows from Theorem 1 and Definitions 3 and 4.

The present experiment was designed to test Theorem 1 above. It failed to confirm it. What are the implications of this failure for the antecedent postulates and for the consequent corollary?

Of the two postulates doubt is cast on the second rather than the first. For the first postulate may stand regardless of Theorem 1, whereas, if the second one holds, the theorem must follow. If, then, we grant that Postulate 2 cannot be accepted as it stands, we must next consider what revision is called for.

From the most conservative standpoint, the postulate is simply too general in scope. It is possible that *under certain conditions* stimuli to a later response are detached from an earlier one, while in other circumstances, including those of our experiment, they are not. What are the chief differences between our experimental conditions and those of the typical problem situations which Guthrie describes? One difference is in the temporal succession of responses. In our experiment the two responses to BG were separated by two other responses to other stimuli, whereas in typical trial-and-error behavior one movement follows on the heels of another. Possibly the dissociation of $S–R_1$ called for by the postulate depends on the closeness with which it is followed by $S–R_2$. Again, it is possible that this dependence itself may be derived from another requirement; namely, R_1, to be inhibited, must be at least incipiently aroused at the time R_2 is made. In our experiment, it will be recalled, the entire difference in favor of the first response was confined to those *S*s who did not recognize the stimulus on its second exposure and who were therefore less likely than the others to repeat the first response.

Another possible limitation to the validity of Postulate 2 lies in the discriminability of the responses involved. It was suggested above that the failure of our experiment to produce inhibition of R_1 by R_2 might be due to the equivalence of the two responses. If two more sharply differentiated reactions had been chosen one might have displaced the other. The grounds for such a deduction must be sought outside of Guthrie's theory. They may be found in the principle of response generalization previously cited. This principle represents one attempt to explain how organisms appear to learn acts rather than specific movements. In the article referred to (**6**), Guthrie came to grips with the same problem. He held that the variety of movements by which an act may be performed is to be understood in terms of the variety of stimulus combinations to which particular movements have been conditioned. Yet his whole argument implies that when a stimulus element enters into a new combination and arouses a new response it thereby becomes detached from any previously associated reaction. There is nothing to suggest that the dissociation is more or less complete according to whether the new response is more or less dissimilar to the old. Guthrie's all-or-nothing conception of conditioning is thus directly opposed to

the concept of response generalization. Here is a nice problem for the experimentalist.[4]

Let us now examine the implications of our negative findings for Corollary 1. As stated above, the corollary affirms that a recurring problem will condition the successful response. It is obvious that by refuting its antecedent theorem we have not refuted the corollary itself. All we have done is to obstruct this particular mode of deriving it. But Guthrie's theory is, in our opinion, the most successful attempt to explain selective learning solely on the basis of associative principles. To what extent are we now forced to abandon Guthrie's approach to the problem of selection and with it, perhaps, the associative approach in general?

The answer depends to a large extent on the final status of Postulate 2. If under certain conditions the postulate stands, in which case Theorem 1 automatically follows, and if those conditions are found to characterize typical problem situations, the corollary can still be derived. Meanwhile it is worth noting that in typical problem situations the principles invoked by Guthrie should operate much more effectively than in the present experiment. In this study, by concentrating on one specific stimulus, BG, buried in the midst of a sequence, we left both responses subject to a large amount of dissociation by the remainder of the series. Besides the colors the *S*s were reacting to other stimuli which persisted throughout, e.g., the panel, the row of pushbuttons, the task set by the instructions. To all of these the final response of the series should, according to Guthrie, remain conditioned. There is some empirical evidence that, in a similar situation, it does. When *S*s are allowed to recall nonsense syllables in any order and recalls are plotted against serial position, the curve rises steeply toward the end of the list (**16, 20**). It even reminds one of the gradients of reinforcement published by other writers (**11, 15, 19**). We suggest that this similarity is more than coincidental. Welch and Burnett (**20**) gave their *S*s a task quite analogous, in certain essential respects, to that which Thorndike—and Guthrie—gave their cats.

Summary

An experiment was designed to test Guthrie's assertion that when a series of responses is made to a stimulus only the final response remains conditioned to it. *S*s were shown a series of colors and were given the task of finding the correct button in response to each. One

[4] In an experiment on transfer of training using lists of paired associates, Bruce (**2**) found that when the stimulus syllables of the two lists were identical there was some negative transfer if the response syllables were different, none if they were similar. The same technique expanded to test retroactive inhibition under these conditions should yield data bearing on the point at issue.

color was shown twice paired with different buttons. In a recall test the stimuli were presented in changed order to determine whether the first or second response to the repeated color was more likely to occur.

The first response was more frequent than the second. For the group as a whole the difference was not statistically dependable; for that portion of the group who did not recognize the color on its second appearance the difference was significant at the 5 percent level of confidence.

It was shown that the results could not be attributed to: (1) intra-serial position, (2) absence of learning, (3) overlearning. An interpretation was suggested in terms of response generalization.

It was pointed out that these findings challenge the general validity of Guthrie's proposition and call for an examination of the conditions under which it does or does not hold good.

(Manuscript received February 12, 1944)

References

1. Berry, R. N., Verplanck, W. S., & Graham, C. H. The reversal of discrimination in a simple running habit. *J. exp. Psychol.*, 1943, **32**, 325–334.
2. Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343–361.
3. Bunch, M. E. Transfer of training in the mastery of an antagonistic habit after varying intervals of time. *J. comp. Psychol.*, 1939, **28**, 189–200.
4. Calkins, M. W. Association. *Psychol. Rev. Monogr. Sup.*, 1896, **1**, No. 2.
5. Fritz, M. F. Long time training of white rats on antagonistic visual habits. *J. comp. Psychol.*, 1930, **11**, 171–184.
6. Guthrie, E. R. Association and the law of effect. *Psychol. Rev.*, 1940, **47**, 127–148.
7. Guthrie, E. R. *The psychology of human conflict.* N. Y.: Harper, 1938.
8. Guthrie, E. R. *The psychology of learning.* N. Y.: Harper, 1935.
9. Harlow, H. F. Studies in discrimination learning by monkeys: I. The learning of discrimination series and the reversal of discrimination series. *J. gen. Psychol.*, 1944, **30**, 3–12.
10. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* N. Y.: Appleton-Century, 1940.
11. Hull, C. L. *Principles of behavior.* N. Y.: Appleton-Century, 1943.
12. Kellogg, W. N., & Wolf, I. S. The nature of the response retained after several varieties of conditioning in the same subjects. *J. exp. Psychol.*, 1939, **24**, 366–383.
13. McClelland, D. C. Studies in serial verbal discrimination learning: IV. Habit reversal after two degrees of learning. *J. exp. Psychol.*, 1943, **33**, 457–470.
14. McGeoch, J. A. *The psychology of human learning.* N. Y.: Longmans, 1942.
15. Muenzinger, K. F., Dove, C. C., & Bernstone, R. H. Serial learning: II. The bi-directional goal gradient in the endless maze. *J. genet. Psychol.*, 1937, **50**, 229–241.
16. Raffel, G. Two determinants of the effect of primacy. *Amer. J. Psychol.*, 1936, **48**, 654–657.
17. Seward, J. P. An experimental study of Guthrie's theory of reinforcement. *J. exp. Psychol.*, 1942, **30**, 247–256.
18. Seward, J. P. Reinforcement in terms of association. *Psychol. Rev.*, 1943, **50**, 187–202.
19. Thorndike, E. L. *The psychology of wants, interests, and attitudes.* N. Y.: Appleton-Century, 1935.
20. Welch, G. B., & Burnett, C. T. Is primacy a factor in association formation? *Amer. J. Psychol.*, 1924, **35**, 396–401.
21. Williams, S. B. Reversal learning after two degrees of training. *J. comp. Psychol.*, 1942, **34**, 353–360.

THE PROBLEM OF STIMULATION DEAFNESS.*
I. COCHLEAR IMPAIRMENT AS A FUNCTION OF TONAL FREQUENCY

BY ERNEST GLEN WEVER AND KENDON R. SMITH

Princeton University

The problem of stimulation deafness arose in the field of clinical otology in consideration of the effects upon the ear of exposure to intense sounds. Investigation in this field soon turned to laboratory animals because with them the traumatic changes may be produced deliberately and under experimental control. For a review of work on this problem up to 1935 the reader is referred to an article by Kemp (1).

The present experiment is one of a series on the problem of stimulation deafness, with emphasis upon quantitative measurements of the exposure tone and the resulting changes of sensitivity, and with consideration also of the nature of individual differences in susceptibility to injury. This report deals with the relation to tonal frequency.

Three stimulating frequencies were used, 300, 1000, and 5000 cycles per sec., and their effects upon sensitivity studied for various tones throughout the auditory range from 100 to 10,000 cycles. Guinea pigs were used as subjects, and the electrical response of the cochlea was employed to indicate the action of the acoustic mechanism.

Procedure

The animals were anesthetized with Urethane, and after the auditory bulla was opened a silver foil electrode was placed in contact with the round window membrane. An inactive electrode was inserted beneath the skin at the site of the incision. The cochlear potentials were led through an amplifier to a General Radio Type 736-A wave analyzer, used as a selective voltmeter.

The stimuli were pure tones from a loudspeaker. The speaker, of a type designed to operate into a horn, was coupled to a tube of one in. diameter, and this tube was stepped down to a glass tube small enough to fit into the external meatus of the ear, where it was securely tied. This 'closed tube' system produced sounds of high intensity with minimum distortion. It had one drawback, however: acoustic damping was small, and variations in the closeness of fit between tube and meatus introduced significant variations in output. Hence it was necessary to calibrate the system for each ear studied.

The sensitivity of the ear was first determined for each of 13 or more tones throughout the range from 100 to 10,000 cycles, by adjusting the intensity of each tone until a standard response (usually of 10 microvolts) was obtained. Thereupon the ear was injured by exposing it to a tone of high intensity for four min. After this exposure, the sensitivity was again determined with the same range of tones as before, in order to assess the damage that had been done. Altogether,

* This investigation was supported by a grant from the John and Mary R. Markle Fund.

31 ears were studied: 11 ears with an exposure tone of 300 cycles, 8 ears with 1000 cycles, and 12 ears with 5000 cycles. The sound pressure was 1000 dynes per sq. cm. for each of these exposure tones.

Results

After exposure to a tone of the extreme intensity employed in this experiment, all animals showed a loss of sensitivity throughout the auditory range. However, the amount of the impairment varied greatly in different ears. Fig. 1 presents some of the results obtained with an exposure tone of 300 cycles.

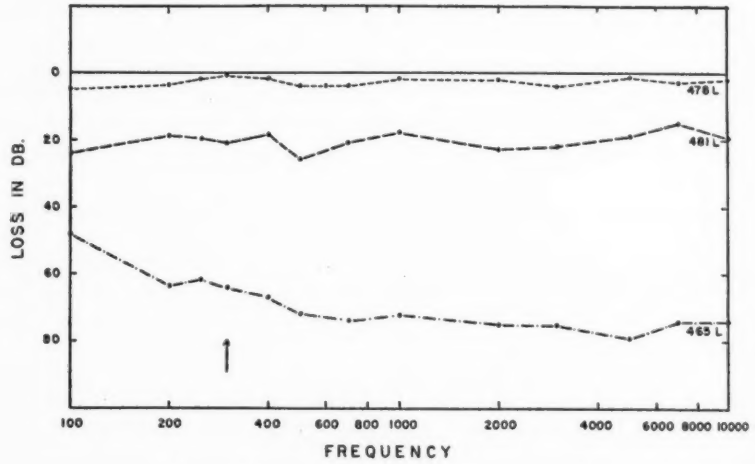


Fig. 1. Injury resulting from exposure to 300 cycles.

In this figure the abscissa represents tonal frequency, and the ordinate shows loss of cochlear response in decibels. The line for zero loss indicates the normal sensitivity initially observed, and the distance of a given curve below this line represents the degree of impairment suffered in consequence of the exposure. Curves are given for three different ears, the numbers of which are shown on the right. These curves were chosen to indicate the range of individual differences found. The uppermost curve, for ear No. 478 L, represents the smallest impairment found among the 11 ears of this group, while the lowermost curve represents the largest impairment in this group. The remaining curve shows an intermediate degree of impairment, a degree more typical of the group.

The range of the impairments is striking. The losses extend from an average of 2.2 db. for No. 478 L to the enormous amount indicated for No. 465 L, which averages 70 db. and in the upper range ap-

proaches 80 db. In the more typical case of No. 481 L the losses are of the order of 20 db.

The impairment is fairly uniform throughout the range studied for the two upper curves, but for the lower curve the effect increases with frequency up to 5000 cycles. It is especially to be noted that there is no sign of a selective effect in the region of the exposure tone.

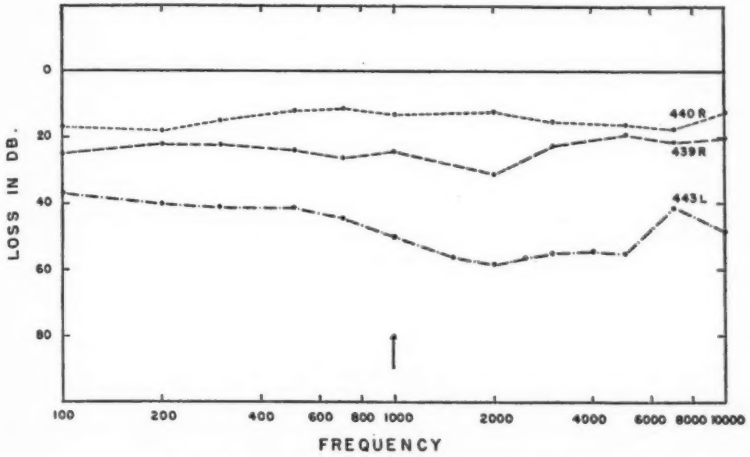


FIG. 2. Injury resulting from exposure to 1000 cycles.

Fig. 2 shows similar results obtained with an exposure tone of 1000 cycles. Again the uppermost curve represents the smallest impairment, the lowermost curve the largest impairment, and the middle curve a more typical amount. As before, the losses are general, and show no specific relation to the frequency of the exposure tone. The range of individual variations is somewhat less than that of the first group, but the average impairment is similar in amount. The lowermost curve shows the most marked effects in the region of the medium high frequencies.

Fig. 3 shows results for an injury tone of 5000 cycles, again for three ears which represent the range of impairments found. The individual variations are still smaller for this group than for 1000 cycles. For the lowermost curve the effects are noticeably less in the high frequency region than for the two lower exposure tones.

A comparison of the effects obtained for the three exposure tones is facilitated by Fig. 4, which contains curves representing averages for the three groups. Here the labels on the curves designate the exposure tone used. The number of animals in each group is limited, and as already indicated the range of individual differences is large,

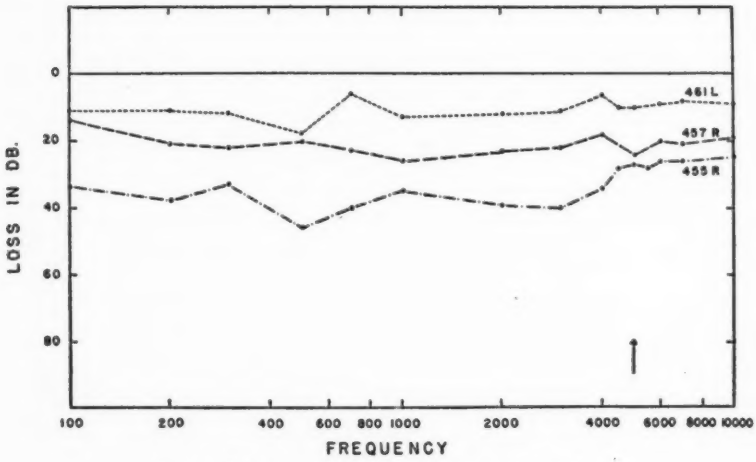


FIG. 3. Injury resulting from exposure to 5000 cycles.

so that too great emphasis should not be placed upon minor variations in these curves. However, it is clear that in general the curves are much alike. There is perhaps a significantly greater effect for 300 cycles in the high-frequency range, but other differences among the curves do not appear significant.

Since the curves for the three groups are similar, and the losses of sensitivity are not specific in respect to frequency, it is permis-

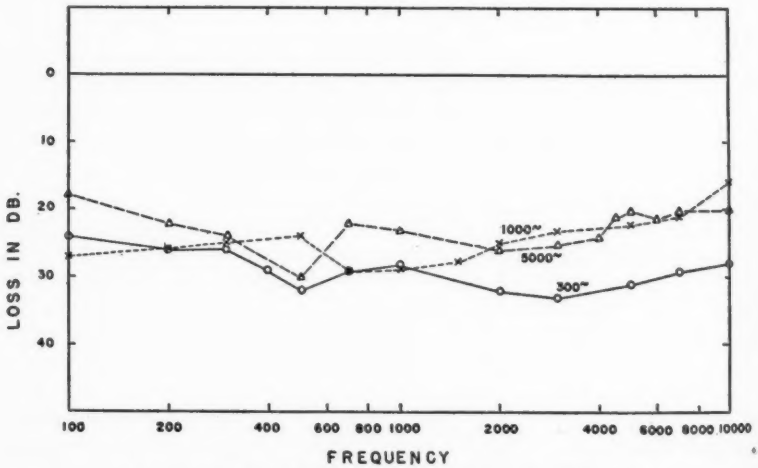


FIG. 4. Average results of injury for three groups of guinea pigs.

sible to combine the observations for the various test frequencies to obtain an overall measure of the effects of injury. The mean loss was computed for each ear by averaging the changes over the frequency range. From these values the group means were obtained, and are given in Table I, together with the mean variations from the mean.

TABLE I

Exposure Frequency	Number of Ears	Average Loss (db.)	Mean Variation (db.)
300	11	28.9	15.7
1000	8	23.4	8.3
5000	12	23.0	5.5

Here it appears that the mean loss was the same for the 1000 and 5000 cycle groups, while it was about six db. greater for the 300 cycle group. However, since six db. is small with respect to the mean variations, this difference can hardly be regarded as significant. The mean variation is greatest for the 300 cycle group, and progressively smaller for the other two groups.

Discussion

In his review of experiments in this field, Kemp noted that clear losses of sensitivity had been obtained on stimulation with tones only within the frequency region of 1000 to 3000 cycles. That range is extended by the results reported here, and by other observations made in this laboratory. Our experiments show that injury may result from stimulation with any tone from 100 to 10,000 cycles, if the intensity is sufficiently great. It seems that the limits in the earlier experiments arose from limitations of the means employed for generating the sounds.

It would appear reasonable that the degree of injury suffered in overstimulation is related to the sensitivity of the ear to the exposure tone. On the other hand, this relation need not be a simple one.

In the present experiments our measure of the sensitivity is the magnitude of cochlear potentials, and the intimacy of the relation that can be shown between sensitivity and injury will depend upon the place of origin of the injury with respect to the origin of the cochlear response. If the injury process concerned some more peripheral structure, like the drum or ossicles, it could easily happen that tones different in frequency but equal as regards physical intensity and aural sensitivity could produce widely different degrees of injury. Equal over-all sensitivity would not insure that the drum or other peripheral structure would be equally acted upon. A closer relation would be expected if the injury arose late in the series of events from

stimulus to cochlear response. We know in the present instance, from histological study, that it is the organ of Corti that suffers disruption from overstimulation, and many considerations point to the sensory cells of this organ as the source of cochlear potentials (2). This close structural association supports our expectation of a correlation between sensitivity and injury.

In examining the results for evidence of this correlation we need to consider both the group data, where the effects of three different frequencies are compared, and the ears within each group, where individual variations appear for single frequencies.

Typically, the sensitivity function for the guinea pig ear shows, at low levels of stimulation, a maximum of sensitivity in the region of 600 cycles, then a marked decline, and then a secondary maximum around 5000 cycles. At high levels, however, the function flattens, and the sensitivity becomes remarkably uniform throughout the range from 200 to 10,000 cycles. At the three exposure frequencies used in the present experiment the sensitivity on the average is nearly the same, at high and low levels alike. Therefore the generally similar injury obtained for the three groups is the expected result.

On the other hand, a close correlation is not found between injury and sensitivity for the individual ears. The rank-order correlation coefficients are 0.45 for the 300 cycle group, 0.02 for the 1000 cycle group, and 0.05 for the 5000 cycle group. Only in the first of these groups is there even a moderate tendency for the more sensitive ears to suffer the larger injuries.

This result is surprising, for at least a small positive correlation ought to appear if the sensitivity reflects the efficiency of transmission of the sound energy to the inner ear. Perhaps this latter assumption is unjustified, and the transmission, from the outer ear up to the locus of the injury process, is highly uniform in normal guinea pig ears. The variations shown in the sensitivity curves then would reflect inner ear characteristics independent of the injury process, perhaps events beyond that process in the series of inner ear actions. Specifically, we can imagine that the injury is determined by the mobility of the basilar membrane, while the sensitivity is a function of such things as the number, distribution, and flexibility of the hair cells of the organ of Corti. The results reported here are insufficient to decide this matter, and it will need to remain open until larger groups are studied, and perhaps other species of animals with different forms of sensitivity curves are investigated.

Summary

Three groups of guinea pigs were exposed to intense tones with frequencies of 300, 1000, or 5000 cycles, and the resulting injuries

were studied in terms of the electrical responses of the cochlea. Widespread impairments were found, with no specific relation to the exposure frequency. Considerable individual differences were observed, particularly for 300 cycles. Though the average impairment was similar in amount for the three exposure groups, the individual ears showed no clear relation between degree of injury and sensitivity to the exposure tone.

(Manuscript received February 16, 1944)

References

1. Kemp, E. H. A critical review of experiments on the problem of stimulation deafness. *Psychol. Bull.*, 1935, **32**, 325–342.
2. Wever, E. G., & Lawrence, M. Tonal interference in relation to cochlear injury. *J. exp. Psychol.*, 1941, **29**, 283–295.

AN EXPERIMENTAL AND STATISTICAL STUDY OF OLFACTORY PREFERENCES [1]

BY MARGARET DAVIES EYSENCK

London, England

Writing in 1939, Hunt (1) pointed out that "The study of what stimuli are pleasant or unpleasant no longer commands the interest it once did. . . . The large individual differences exhibited have precluded the possibility of any exact prediction of the hedonic value of stimuli for specific individuals." Interest has recently revived, as shown by the researches summarized by Peters (2) and Wiley (3), and new statistical methods have been applied to this field in an attempt to assess the relative contribution of general, type, specific and error factors (4, 5).

In the field of olfactory hedonic judgments a dispute has arisen regarding the comparative importance of these factors, which seems to be due primarily to a misunderstanding of the statistical issues, and it seemed worthwhile to try to settle the controversy by means of a new experiment and a re-analysis of the previous data.

Beebe-Center (6), correlating the rankings of 14 smells by eight *S*s, analyzed the resulting table in terms of Spearman's method of tetrad differences, and found "that the table satisfies the tetrad criterion, and consequently we may infer that the olfactory substances may be considered to have a general hedonic value which may be called $h_{olf.}$"

Stephenson (7), repeating this investigation with 14 *S*s and 31 smells, denied "that the person with the highest g saturation . . . has the most ability," i.e., had "the keenest hedonic susceptibilities—on the contrary he is the most ordinary person, the most typical." He also claimed to have shown the existence of a type factor in hedonic judgments of olfactory stimuli in contradiction to Beebe-Center's claim that a general factor only was present.

In my own investigation, 14 *S*s, eight adult men and six women, mostly University students or persons of the same cultural level, were asked to arrange 31 smells (some of which were pleasant, others unpleasant) in accordance with the following frequency distribution. This distribution, although it departs a good deal from the normal probability curve, was used to make the results as nearly as possible comparable with those of Stephenson in his experiment.

[1] I am indebted to Professor C. Burt for his invaluable help in this investigation.

Score:	10	9	8	7	6	5	4	3	2	1	0
Frequency:	1	2	2	3	4	7	4	3	2	2	1

Each stimulus was presented in a small, corked, dark green bottle and labelled only by a letter. Care was taken to ensure so far as possible that olfactory fatigue was avoided by the order of presentation of stimuli.[2] The usual procedure was followed of having the *S*s make first of all a rough grouping of the smells into 'definitely pleasant,' 'average' and 'definitely unpleasant,' and then rank the stimuli, giving the best liked odor a score of 10, the next two best liked 9, and so on, down to the least liked (or most disliked) a score of 0.

Correlations between the persons were worked out and factor saturations obtained by the least squares method described by Burt (4). After removal of the first factor, the residual coefficients were calculated. None of the residuals was statistically significant as judged by the usual criterion (two times the standard error of the difference given by Fisher's formula, the standard error being 0.267). The first or general factor accounted for approximately 58 percent of the variance, while the second factor contributed only seven percent.

This predominance in importance of the first or general factor accords well with the conclusion the writer (9) drew from a re-analysis of all researches reported to date in which correlations between persons had been used, namely that "In 44 out of the 48 researches examined the first factor will account for all the variance and covariance within the limits imposed by the probable error."

As both Beebe-Center and Stephenson give their intercorrelations, I applied the method of least squares to their tables, and give below the variances found for the first and second factors in their tables, together with those from my own.

	Factor I	Factor II
Beebe-Center	0.571	0.084
Stephenson	0.499	0.088
Eysenck	0.582	0.068

Although Beebe-Center's stimuli included no definitely unpleasant odors, as Stephenson's and my own did, all three sets of results would appear to be in close agreement. The interpretation of the first factor presents little difficulty. Beebe-Center speaks of 'aesthetic sensitivity': "Given a set of objects to which there have been assigned general hedonic values (i.e., values more or less common to all observers but independent of the particular peculiarities of the indi-

[2] As Kniep et al. (8) have shown by experimental means, "an odor has a general affective value which to a large extent is relatively independent of its position in a series."

vidual observer) the aesthetic sensitivity of an individual to these objects is a function of the correlation between his order and the general order." Thus, a person with a high saturation coefficient for the first factor is a person whose order approximates most closely to the average or general order.

It is interesting to speculate on the reasons for this general order. Two possible causes which suggest themselves are (1) innate, biological factors, such as dislike of nocive stimuli, and (2) associations acquired due to social learning. In the related field of color preferences H. J. Eysenck (**10**) has shown that white *S*s and a variety of colored *S*s have a very high agreement in their average ranking ($r = 0.96$), although their associations to the colors would be profoundly different, and he tried to relate these preferences to physical factors. The sense of smell being more primitive than that of sight, it would appear even more likely that the general order is based on biological factors. The probability of this being so is strengthened by the fact that the average intercorrelation on hedonic judgments of smell is considerably higher than that on hedonic judgments of color, the reason presumably being the greater strength of the biological influence and the correspondingly weaker effect of individual associative factors.[3]

It is perhaps important to mention the necessity of distinguishing between the sort of general factor established here and a 'unitary' factor. It may well be that the latter type of factor exists for olfactory judgments, but to ascertain whether it does or not we would have to give a number of separate tests of olfactory stimuli to the same subjects, get their factor saturations in each test, and then intercorrelate their saturations. If from this table of correlations a general factor emerged, then we should have shown that a 'unitary' factor of aesthetic sensitivity to smell is present. In other words, we must test whether or not a person who judges odors most in accord with the universal order in one sub-test tends to do the same in other sub-tests. In the case of æsthetic judgments on pictorial art such a unitary factor was clearly demonstrated (**12**), while in the case of judgments on the 'goodness' of various types of humorous material, although for each separate test there was a large measure of agreement among the *S*s, no such unitary factor was apparent (**13**).

Beebe-Center, as previously mentioned, found that his table satisfied the tetrad criterion, and did not work out any further factors.

[3] Kniep et al. (**11**) reached a similar conclusion when they found high agreement between the judgments of children and adults, and argued: "This degree of constancy may be attributed to physiological mechanisms which are relatively mature at the age of 7." However, it may be doubted whether the high measure of agreement between adults and children precludes the possibility that it is based on associations due to social learning which might also be expected to be comparatively fixed by the age of 7.

Stephenson explains his second factor, whose statistical significance he does not test, by unusual and divergent responses to the tar and resinous odors. However, when Stephenson's first factor is extracted by the method of least squares, only one of the 91 residuals is more than twice the standard error, and in a table of this size it would be surprising if chance did not produce at least one such deviation. Accordingly, we are unable to accept the significance of his second factor.

None of the residuals on which my second factor is based was statistically significant. However, as a number of the *S*s who took part in this investigation coöperated later in a research conducted by H. J. Eysenck (**14**) into aesthetic judgments of landscapes, portraits, statues, etc., it seemed interesting to discover whether or not the second factor in this present research might not have some relation to the strong bipolar brought out in the subsequent investigation into visual preferences. In spite of the small number of *S*s a statistically significant correlation was found, showing that those who preferred the bright, highly colored, modern type of art also tended to show a preference for the stronger, more pungent odors. This result would lend a certain amount of support to the interpretation of this bipolar factor in terms of Hornbostel's general 'brightness factor,' but, of course, the small number of *S*s used and the fact that the type factor in the present study was based on non-significant residuals makes the repetition of this experiment desirable before any definite conclusions can be drawn.

Below is a list of the odors used as stimuli, in the weighted [4] order of preference:

1. Verbena, 2. Vanilla, 3. Cinnamon, 4. Eau de Cologne, 5. Amyl Acetate (Pear Drops), 6. Lavender Oil, 7. Aniseed, 8. Camphor, 9. Cloves, 10. Eucalyptus, 11. Pennyroyal, 12, Carbolic Acid, 13. Creosol, 14. Iodoform, 15. Beeswax, 16. Pine, 17. Coffee, 18. Rhubarb, 19. Tea, 20. Lactic Acid, 21. Parsley, 22. Paraldehyde, 23. Cod Liver Oil, 24. Sage, 25. Benzene, 26. Acetone, 27. Ether, 28. Valerian, 29. Asafoetida, 30. Amyl Nitrate, 31. Pyridinal.

Although the above list is based on the judgments of only 14 *S*s, the agreement among their rankings is so high that it can be shown that their average order of preference for these smells correlates with the order for the entire population of which these 14 are a true sample to the extent of 0.975 (**15**).

Using Zwaardemaker's classification, and taking the weighted order of preference for the smells under each of the nine headings, the following average positions appear:

[4] Each score was weighted, using the formula $W_u = \frac{r_{ug}}{1 - r^2_{ug}}$, where r_{ug} is the Factor I saturation for any person 'u' and W_u his appropriate weight.

Aromatic	7.5
Balsamic	10.5
Etherial	15.7
Empyreumatic	17.4
Repulsive	18
Amber-Musk	22.5
Hircine	24
Alliaceous	25.7
Nauseous	26

This seems to confirm the usual view about the relative pleasantness and unpleasantness of these groups.

Stephenson (**7**) had remarked that "The turning points from pleasant to unpleasant do not appear at the same smells for all the persons." A similar conclusion had been reached previously by Kniep et al. (**8**). The *S*s who took part in the present experiment were interrogated about this, and a like divergence was discovered. One person said 21 of the 31 odors were unpleasant, while two found that only five were definitely unpleasant. The women, on the average, disliked 9.5 and the men 11.75. The correlation between the Factor I saturation and the fewness of the odors disliked was calculated. The correlation was only 0.39 ± 0.24 S.E., but perhaps it indicated that 'pleasure in smelling' tends to be coupled with a high general factor saturation, or capacity to approximate the average judgment in ranking.

The average intercorrelation was 0.573 in this experiment; the men agreed among themselves to the extent of 0.511; the women among themselves to the extent of 0.655; the average order for the men correlated with that for the women to the extent of 0.895. The theoretical correlation between the average rankings of the sexes is 0.906 (the product of their respective correlations with the true average order (**15**)). From the fact that the real correlation between the sexes does not differ significantly from the theoretical, we can conclude that no sex differences in olfactory preferences are apparent in our *S*s. The average general factor saturation for the men is 0.717 and for the women 0.811.

Widely divergent judgments by certain individuals were observed for a few of the stimuli, namely Cinnamon, Pine, Pennyroyal (similar to Peppermint), Creosol, Lavender, Cloves, Vanilla, Eucalyptus, Iodoform. There was no evidence in the present research that divergent judgments on any of these odors tended to establish 'types';[5] it is possible that further work might indicate that some at least of these smells give rise to 'specific' factors, in which case the indi-

[5] The best method of establishing 'type' factors in a field where a strong general factor is present appears to be that first suggested by H. J. Eysenck (**14**), who eliminated the general factor experimentally rather than statistically by a suitable selection of stimuli, and was thus able to prove the existence of significant bipolar or type factors in the field of visual preference rankings.

vidual's abnormal judgment on the stimuli would be persistent; [6] the third possibility is that what are called 'error factors' or 'chance factors' give a complete explanation of these unusual judgments.

Summary and Conclusions

Fourteen *S*s were asked to rank 31 olfactory stimuli, allotting them scores in accordance with a prescribed frequency distribution. The intercorrelations of their rankings were factor-analyzed by the method of least squares.

The tables of intercorrelations found by Beebe-Center and Stephenson from similar experiments were re-analyzed by the same method, and the opposing interpretations given by these earlier workers in the field were discussed on the basis of the statistical results.

The following conclusions were reached:

1. There is evidence for the existence of a general factor in olfactory judgments, accounting for approximately 58 percent of the variance.

2. In none of the three researches was conclusive proof found for the view that any 'types' exist in olfactory judgments.

3. No sex differences in olfactory preferences were discovered.

4. The average preferences for smells in the various olfactory groups substantiate the usual view: those in the Aromatic and Balsamic categories being the most pleasant and those in the Nauseous the least pleasant.

5. Agreement between data derived from olfactory and visual preference judgments gives a certain amount of support to the view that there exists an intersensory factor of 'brightness.'

(Manuscript received February 14, 1944)

References

1. Hunt, W. A. A critical review of current approaches to affectivity. *Psychol. Bull.*, 1939, **36**, 807–828.
2. Peters, H. N. The experimental study of aesthetic judgments. *Psychol. Bull.*, 1942, **39**, 273–305.
3. Wiley, L. N. A resume of major experiments in abstract design ending with the year 1941. *Psychol. Bull.*, 1941, **38**, 976–990.
4. Burt, C. *The Factors of the mind.* London: Univ. of London Press, 1940.
5. Eysenck, H. J. The experimental study of the 'good Gestalt'—a new approach. *Psychol. Rev.*, 1942, **49**, 344–364.
6. Beebe-Center, J. B. *Pleasantness and unpleasantness.* New York: Century, 1933.
7. Stephenson, W. The inverted factor technique. *Brit. J. Psychol.*, 1936, **26**, 344–361.
8. Kneip, E. H., Morgan, W. L., & Young, P. T. Individual differences in affective reaction to odors. *Amer. J. Psychol.*, 1931, **43**, 406–414.

[6] The probability that such specific factors exist in this field is indicated by the work of P. T. Young (**16**, **17**).

9. DAVIES, M. The general factor in correlations between persons. *Brit. J. Psychol.*, 1939, **29**, 404–421.
10. EYSENCK, H. J. A critical and experimental study of colour preferences. *Amer. J. Psychol.*, 1941, **54**, 385–394.
11. KNIEP, E. H., MORGAN, W. L., & YOUNG, P. T. Relation between age and affective reaction to odors. *Amer. J. Psychol.*, 1941, **43**, 414–421.
12. EYSENCK, H. J. The general factor in aesthetic judgments. *Brit. J. Psychol.*, 1940, 31, 94–102.
13. EYSENCK, H. J. An experimental analysis of five tests of 'appreciation of humor.' *Educ. & Psychol. Measurement*, 1943, **3**, 191–214.
14. EYSENCK, H. J. 'Type'-factors in aesthetic judgments. *Brit. J. Psychol.*, 1941, **31**, 262–270.
15. EYSENCK, H. J. The validity of judgments as a function of the number of judges. *J. exp. Psychol.*, 1939, **25**, 650–654.
16. YOUNG, P. T. Constancy of affective judgments to odors. *J. exp. Psychol.*, 1923, **6**, 182–191.
17. YOUNG, P. T. The scale of values method. *Amer. J. Psychol.*, 1930, **42**, 17–27.

A DUAL SELECTIVE AMPLIFIER

BY HILDING B. CARLSON AND ALBERT JAMES EBEL

University of Illinois

To gain greater standardization than is possible by manual means, a dual selective amplifier was designed and constructed with an especially cut set of test records for the administration of a blinker modification (1) of Thurstone's 1943 Code Aptitude Test (2). This set-up enabled the experimenters to give instructions auditorially concerning Morse Code signals presented visually, the time intervals between instructions and code signals, as well as the speed at which the signals are sent, being held constant from one testing period to another.

The basic principle of the dual selective amplifier is not new. However, since its applications to psychological research and testing are broader than for the particular test for which it was constructed, its design is given here. The particular constants used for this unit are given in Diagram 1, with a description of the essential parts included in Table I. Both diagram and table are self-explanatory,

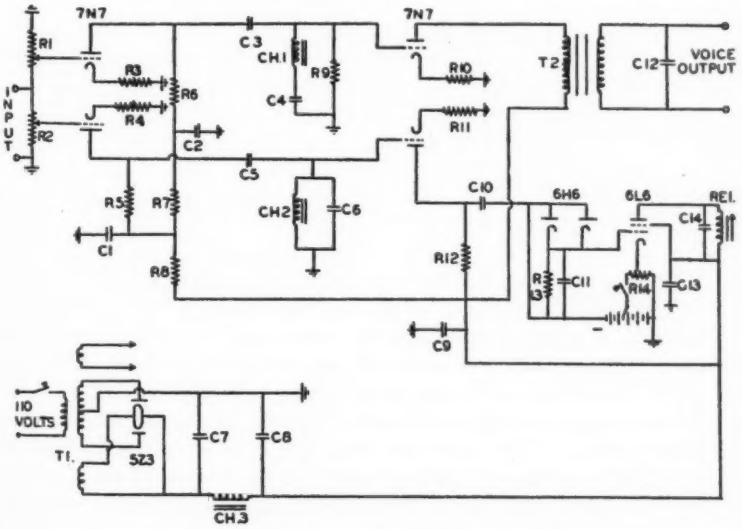


DIAGRAM 1. Wiring diagram for dual selective amplifier.

TABLE I

PARTS LIST FOR DUAL SELECTIVE AMPLIFIER

Item	Description
Resistors	
R1	500,000 ohm voltage divider with uniform resistance change
R2	1,500,000 ohm voltage divider with uniform resistance change
R14	25,000 ohm voltage divider with uniform resistance change
R3, R4	500 ohms, 2 watts
R5	100,000 ohms, 2 watts
R6	30,000 ohms, 2 watts
R7, 12	25,000 ohms, 2 watts
R8	10,000 ohms, 2 watts
R9	250,000 ohms, 2 wotts
R10	1,000 ohms, 2 watts
R11	1,500 ohms, 2 watts
R13	35,000 ohms, 2 watts
Condensers	
C1, 5, 12, 13	0.5 Mfd. paper
C2, 11	0.1 Mfd. paper
C3	0.001 Mfd. paper
C4, 6, 9	1.0 Mfd. paper
C10, 14	2.0 Mfd. paper
C7, 8	8.0 Mfd. dry electrolytic
Choke coils	
CH1, 2	Type C2304, 8 Henries, 150 Ma., 65 ohms
CH3	Type C1721, 15 Henries, 150 Ma., 65 ohms
Transformers	
T1	Universal type; 700 plate volts at 150 Ma.; Filament 1, 5 volts at 3A.; Filament 2, 6.3 volts at 4.7A
T2	Type A3304 output with 500 ohm secondary
Relay key	
RE1	Type MRS291 SPDT 2500 ohms plate relay
Tubes	
7N7	2 dual-triode
6H6	1 diode
6L6	1 beam power
5Z3	1 rectifier
Batteries	
"C"	45 volts
	22½ volts

except for the connections for voltage divider R14, which are attached across the nine volt terminals of the 22½ volt 'C' battery.

Sixteen-inch phonograph records were used in connection with this dual amplifier and were cut and operated at 33⅓ revolutions per min. This was done in order to minimize the number of records needed for the test as well as the necessity for frequent pauses while changing records. They were cut with the instructions recorded at plus 16 decibels for our recording head and the code signals, using a 78-cycle tone produced by a Hewlett-Packard audio oscillator, at plus 13 decibels. The important factor here is that the code signals be recorded at least three decibels below that for the voice. These

relatively low intensities were necessary in order to prevent the speech from operating the relay key and the 78-cycle tone from being heard in the loud speaker.

Normal operation of a telegraph key produces a click at the onset and cessation of each operation under the conditions indicated. In order to prevent the upper frequencies of the clicks from being heard, it was necessary to find some way to eliminate or at least diminish them. While various techniques might have been used, it was accomplished in this instance by a condenser-choke coil hook-up that made the 78-cycle tone continuous throughout the recording, but at so low a level of intensity when the telegraph key was not operated that it is not heard. Diagram 2 gives the wiring plan used.

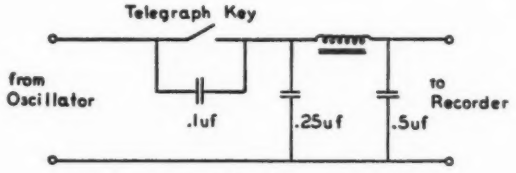


DIAGRAM 2. Hookup for controlling clicks.

It was found desirable to have a woman give the instructions since the lower pitch of a man's voice occasionally operated the relay key during preliminary trials.

In operation, the output of the phonograph pickup is connected to the two grids of the first dual-triode amplifier. The two plates of this amplifier tube are resistance-capacity coupled to the two grids of the second dual-triode amplifier. In the grid circuit of the voice section of this dual-triode is a series resonant circuit that eliminates the signaling tone from this section of the dual amplifier. In the signal tone section of this dual-triode a parallel resonant circuit is used to eliminate all other frequencies except the signaling frequency. Thus the voice and signaling tone are effectively separated. Care must be taken to shield all grid connections in the voice section, as well as the output transformer, in order to prevent cross-talk between the two circuits. The output of the voice section of the second amplifier tube is coupled to the external circuit through an output transformer with a 500-ohm secondary. The output of the signal section of the amplifier is rectified in the 6H6 diode and filtered with condensers across the diode load resistor. The constants of this circuit have been carefully chosen so that the time constant is not longer than the keying period for the dots in the signals. Substitution of other constants may result in improper functioning of the relay key.

The 6L6 relay tube is normally biased so that the relay in the plate circuit does not close. This is approximately cut-off bias. The rectified voltage obtained from the signal amplifier tube and the rectifier tube is used to reduce this bias so that the plate current increases sufficiently to operate the relay. Filter condensers are also used across the relay to remove the A.C. component of the signal voltage which without the condensers tends to make the relay chatter. If too large a condenser is substituted for that indicated, the time constant of the circuit will be such that the relay will not follow the dots in the signals.

When the test records are played, potentiometer R2 is adjusted to control the code signal input, (and potentiometer R1 to control the voice imput). The setting of potentiometer R14, which controls the bias on the 6L6 tube, is very critical, for with insufficient bias the relay key is activated constantly, and with too much bias, it fails to respond at all. At the proper setting, however, the relay key responds to the 78-cycle tone only and thus operates the blinker lights. At the same time the relay key fails to respond to the higher frequencies involved in speech, if the voice section of the dual amplifier does not introduce any distortion into the signal, so that the higher overtones of the signaling tone lie in the acceptance range of the voice amplifier. One method of obviating this difficulty, if it should occur, would be to record at a lower level.

While the dual amplifier introduces some slight distortion of the voice, it is not sufficient to interfere with understanding. The distortion is less when a woman's higher pitched voice is used than for a man's voice.

(Manuscript received February 20, 1944)

References

1. Carlson, Hilding B., & Rich, J. A blinker adaptation of Thurstone's 1943 code aptitude test. *Psychol. Bull.*, 1944, **41**, 322–331.
2. Thurstone, L. L. Report on a code aptitude test. Privately printed. May, 1943.